New-Product Forecasting

New-Product Forecasting

Models and Applications

Edited by
Yoram Wind
Vijay Mahajan
University of Pennsylvania
Richard N. Cardozo
University of Minnesota

LexingtonBooks
D.C. Heath and Company
Lexington, Massachusetts
Toronto

Library of Congress Cataloging in Publication Data
Main entry under title:

New-product forecasting

Includes index.
1. New products—Mathematical models—Addresses, essays, lectures. 2. Product management—Mathematical models—Addresses, essays, lectures. 3. Sales forecasting—Mathematical models—Addresses, essays, lectures. I. Wind, Yoram. II. Mahajan, Vijay. III. Cardozo, Richard N.
HF5415.153.N48 658.5'75'0724 80-8388
ISBN 0-669-04102-5
ISBN 0-669-09806-x Paperbound

Third printing, November 1984

Published simultaneously in Canada

Printed in the United States of America on acid-free paper

International Standard Book Number: 0-669-04102-5 Casebound

International Standard Book Number: 0-669-09806-x Paperbound

Library of Congress Catalog Card Number: 80-8388

Contents

Preface and Acknowledgments

The forecasting of new-product performance is one of the most difficult and critical management tasks. The actual market performance of a new product depends on a large number of factors including not only consumer attraction to and satisfaction with the product, which in turn can vary depending on external conditions (for example, economic, political, and so forth), but also the degree of trade support for the product and competitive reactions to it. Assessing these uncertainties makes the forecasting of new-product performance much more complex than the forecasting of the sales share and profit of established products. Yet, accurate forecasting of new-product performance is a must and constitutes a major input to most corporate and marketing decisions.

Given the importance of accurate forecasting of new-product performance and the intellectual challenge of developing better models, numerous new-product forecasting models have been developed and proposed in the marketing and management-science literature. Yet, despite the continuous attention given to this area, there have not been any major books on the topic that attempt to summarize and evaluate the various models and assess their actual implementation. The purpose of this book is to fill this gap and offer, for the first time in one place, a systematic coverage of new-product forecasting models. The unique characteristics of this book are:

1. It presents, in full, fourteen of the major new-product forecasting models. These include not only the well-known models of Bass, N.W. Ayer, Parfitt and Collins, and others, but also a number of models and approaches that have not appeared in the literature. These include, for example, the POSSE system developed by Green and the commercial pre-test-market simulators of Yankelovich and of Elrick and Lavidge.
2. It offers a classification of new-product forecasting models that is based on the type of data required for implementing the model. The classification divides new-product forecasting models into four classes, corresponding to four major stages in the new-product development process. These four new-product forecasting approaches are based on data from concept-test studies, pre-test-market studies, test market, and early sales. The models presented in each group are those which required for their development the type of data referred to. Obviously,

to the extent that a given model has been developed and its parameters have been estimated, it can be used at other stages in the new-product development process.

3. The discussion of each group of models includes an introductory overview chapter, a full description of three or more illustrative key models and approaches, and a chapter discussing the utilization of these types of models by industry. In most cases, the overview and utilization chapters have been prepared especially for this book. Furthermore, the utilization chapters have been prepared by practitioners who have had firsthand experience in the utilization of these and other models.
4. Given the large number of available new-product forecasting models, the first chapter of the book is designed to help evaluate and select appropriate models. To this end, it presents a framework for classification and evaluation of new-product forecasting models.

This book is designed as a reference book for practitioners concerned with the selection and evaluation of new-product forecasting models. It is our hope that it will also provide a stimulus for further empirical comparisons of the various models and the development of a better understanding of the determinants of new-product success and models for forecasting it.

We are indebted to the thirty-four authors and the following publishers and journals who allowed us to reprint their material: Addison-Wesley, the American Marketing Association (*Journal of Marketing Research, Journal of Marketing*), *Journal of Consumer Research, Harvard Business Review,* and The Institute of Management Sciences (*Management Science*). We would also like to thank Lynda Kenny and Billie Meeks for assisting us in our contacts with authors and publishers and coordinating and typing the manuscript.

Part I
Classification and Evaluation of New-Product Forecasting Models

The large number and variability of current new-product forecasting models raise the question of which is the most appropriate. Answering this question requires familiarity with both new-product forecasting models and the criteria for their evaluation. Chapter 1 proposes a multidimensional scheme to differentiate between and classify new-product forecasting models. This scheme is based on eight sets of characteristics: the purpose of the model, the type of products and services that can be studied, the unit and level of analysis, the model format, the dependent and independent variables, the required data, and the analytical procedures employed.

A set of criteria for the evaluation of the various new-product forecasting models is proposed next. The obvious and most important criterion is the model's actual performance, that is, the accuracy of its forecast. Yet, this criterion is often of limited practical value in evaluating various new-product forecasting models since most models are claimed (by their developers) to provide high-level accuracy. If one assumes that two models can predict equally well, then the question must be: What other criteria can management use in evaluating alternative new-product forecasting models? In addition to the predictive accuracy, chapter 1 suggests that the ability to develop and implement the model and its diagnostic power are the key criteria for evaluating new-product forecasting models.

This part, in addition to presenting an overview of the new-product forecasting models, introduces the issues involved in developing and selecting new-product forecasting models and briefly reviews some key new-product forecasting models.

1 A Framework for Classifying New-Product Forecasting Models

Yoram Wind

The Proposed Framework

The proposed framework for the classification of new-product forecasting models is presented in table 1-1 and is based on eight sets of characteristics—the purpose of the model, the type of products and services that can be studied, the unit and level of analysis, the model format, the dependent and independent variables, the required data, and the analytical procedures employed.

Purpose of the Model

New-product forecasting models vary considerably with respect to their objectives. Some forecast total market demand (that is, ignoring the distinction between trial and repeat purchase); others forecast first purchase (trial) (for example, Eskin 1973); still others forecast also repeat purchases (for example, Claycamp and Liddy 1969). A second objective-related distinction among new-product forecasting models involves the use of the model for predictive purposes only (Ehrenberg 1971) versus its additional utilization as a diagnostic tool that provides insight into the nature of the response function to various marketing variables and strategies (Claycamp and Liddy 1969). In addition, one can distinguish among forecasting models based on the forecast horizon (number of years forecasted) and the forecasting periods (weekly, monthly, quarterly, or annually).

Type of Products and Services Covered

New-product forecasting models can be classified according to the frequency of purchase of the given product or service—frequently purchased items (Claycamp and Liddy 1969) versus infrequently purchased ones, such as durables (Bass 1969; Brown, Buck, and Pyatt 1965; and Ryans 1974).

Reprinted from Yoram Wind, *Product Policy: Concepts, Methods, and Strategy* (Reading, Mass.: Addison-Wesley, 1981), chap. 15. Reprinted with permission.

Table 1–1
A Framework for Classifying New-Product Forecasting Models

Purpose of Model

Forecasting of aggregate market demand versus trial- or repeat-purchase forecast
First-purchase models versus repeat-purchase models
Prediction versus prediction and diagnostics
Forecasting of total sales versus weekly, monthly, quarterly, or annual sales

Type of Products and Services

Frequently purchased versus infrequently purchased (durables...)
New product versus new brand in an established product class
Consumer versus industrial products and services

Unit and Level of Analysis

Unit: Individual versus household
Consumers versus units bought
Level: Market segment versus total market

Format of Model

Diffusion versus adoption/behavioral models
Deterministic versus stochastic

The Dependent Variable

Type of Variables:
Brand choice
(trial) versus frequency of purchase
(repeat) versus quantity purchased
Brand sales versus brand share

Number of Variables:
Single versus two or more

Type of Scale:
Nominal-ordinal interval-ratio-mix

The Independent Variable

Type of Variables:
Time
Marketing variables (price, deals, advertising, etc.)
Customer characteristics—situation-specific (attitude, awareness, etc.) versus general customer characteristics
Competitive activities
External environmental forces

Number of Variables:
Single versus two or more

Type of Scale:
Nominal-ordinal interval-ratio-mix

Required Data

Type of Data
Primary data versus historical data on similar products (analogy)
Consumer response to actual product versus response to concept description
Consumer responses versus sales data
Test market versus pretest market with an existing product (in home-use test)
Panel versus nonpanel data
Data-collection procedure (telephone, mail, personal)
Number of observations required
Time and cost of data collection

Analytical Procedures

Simulation versus analytical procedures
Parameter estimation across product classes versus within a product class
One-step versus multiple-step approach

Whereas this distinction is quite arbitrary—products are bought at different levels ranging all the way from every day to once in a lifetime—it does serve as a useful classification variable since it provides the rationale for focusing on single-purchase behavior versus trial and repeat-purchase forecasts.

The second aspect of type of product involved is the *newness* of the product in question—whether the product presents a completely new product class or a new item in an established product class. In general, the greater the newness of the product, the greater the need to base the forecast

on actual experience with, and use of, the product. If an early reaction to a totally new product is desired (that is, at the concept-testing stage), the research design (and forecasting model used) should ensure that the consumers truly understand the new product. This can be done by visual aids which "educate" the respondents. Finally, one may distinguish between consumer and industrial products and services. Given, however, that with the exception of conjoint-analysis-based forecasting simulators, there has been a lopsided emphasis on forecasts of consumer products, the appropriate distinction should be the potential *applicability* of the model to the forecast of industrial products and services.

Unit and Level of Analysis

New-product forecasting models differ also with respect to the unit of analysis employed by them—whether it is the individual (Blattberg and Golanty 1974) or the household (Lipstein 1970). A second distinction is between a forecast of the number of consumers (either individuals or households) and the number of units bought. This distinction is essential when consumers differ in the level of product use. A meaningful forecast of the demand for a specific airfare for a flight across the North Atlantic, for example, is the number (and share) of *trips,* not the number (or share) of consumers preferring this fare, although the latter information can be important diagnostically.

In addition to the unit of analysis decision, it is important to determine the level of analysis—a market segment or the total market. The selection of the "right" level of analysis (including the level of segmentation) depends on the degree of homogeneity of the given market and the firm's segmentation strategy. From a diagnostic point of view, forecasts by segment are more helpful, and it is often desirable to compare the results one can obtain from a total-market forecast with those derived from an aggregate of the separate segment forecasts.

Format of Model

New-product forecasting models can be classified based on two sets of characteristics concerning their format. First is the diffusion versus the adoption/behavioral nature of the model. Diffusion models are based on the forecasting of a *product's* life-cycle function, whereas adoption models focus on the *individual's* adoption process of moving from a state of unawareness toward the trial of a new product. The distinction between diffusion and adoption was considered at least as early as 1962 by Rogers. Diffusion-type models include a number of distinct models ranging from

exponential models such as the Fourt and Woodlock (1960) model to the logistic (S-shaped) model such as the one used by Griliches (1967) and Mansfield (1961), through an epidemiological diffusion model as described by Bailey (1957) and modified as a new-product forecasting model by Bass (1969), to the reliability engineering diffusion model utilized by Burger (1968, 1972) in his new-product forecasting model.

The adoption/behavioral forecasting models, on the other hand, are based on explicating the adoption process with or without explicit inclusion of marketing variables. The DEMON model (Charnes et al. 1966, 1968) and the BBDO NEWS model, briefly summarized in figure 1-1, are examples of forecasting models of this type.

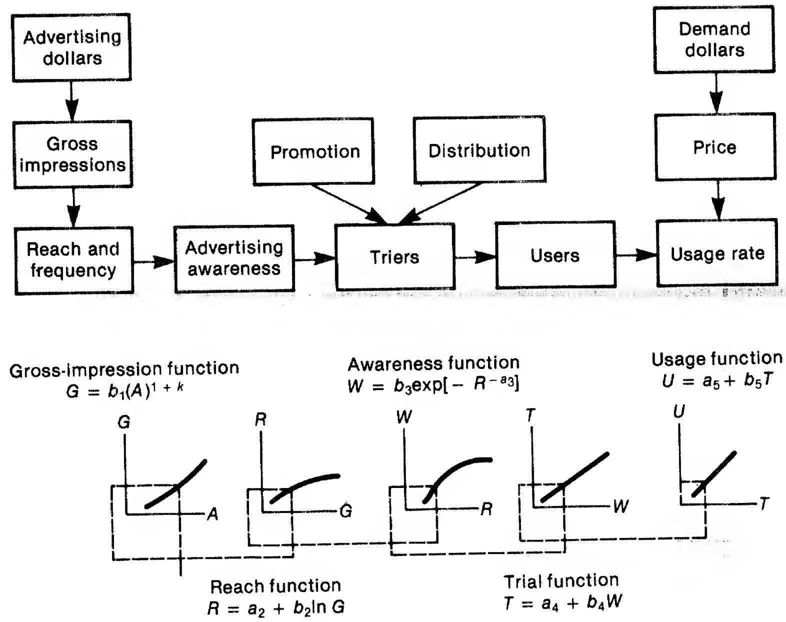
The second classification is based on the deterministic versus stochastic nature of the model. The stochastic nature of a new-product forecasting model (see, for example, Nakanishi 1973) can be expressed in the model's functional format, assumptions (for example, constant versus changing rate of repurchase), and the nature of the output, whether the output of the forecast is presented as a point estimate, as in most models, or as a probability distribution (discrete or continuous) of possible estimates.

Dependent Variable

New-product forecasting models differ considerably in terms of the dependent variable they employ. Whereas some use brand choice (Eskin 1974), others focus on the frequency (Herniter 1974) or the quantity to be purchased (Hamburg and Atkins 1967). Furthermore, while most models focus only on the forecast of a given brand's sales, few employ a brand's market share as the dependent variable (Cima et al. 1973). Ignored here are models which do not use *purchase* (sale) measures, but rather focus entirely on other consumers' responses such as awareness or attitudes. Two exceptions are, however, models which use a measure of consumers' *intentions* to buy a product as the sole dependent variable and the more common case in which a number of dependent variables are used, including a number of purchase-based and non-purchase-based variables. The most common (and managerially critical) of such variables are the following.

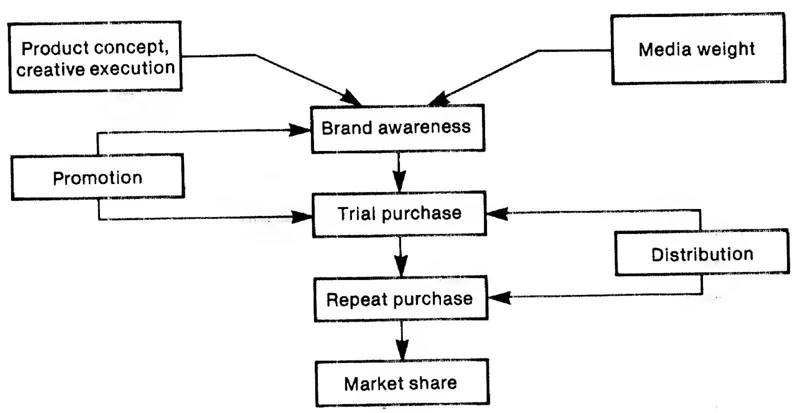
Awareness

This variable comprises the percentage of consumers aware of a brand. This dependent variable measures the effectiveness of the introductory advertising and promotional campaign and often serves as the basis for calculating the percentage of triers.



Source: D.B. Lerner, "Profit Maximization through New Product Market Planning and Control," in F.M. Bass, C.W. King, and E.A. Pessemier (eds.), *Applications of the Sciences in Marketing Management* (New York: Wiley, 1968) pp. 151-167.

Figure 1-1a. The Demon Model



Source: BBDO, Management Science Department, *The Theoretical Basis of NEWS* (New York, 1971).

Figure 1-1b. The NEWS (*N*ew-*P*roduct *E*arly *W*arning *S*ystem) Model

Trial

This variable comprises the percentage of consumers who are aware of the brand and tried it. Trial models often parallel epidemiological models. Yet, in the introduction of a new brand, trial is not always a meaningful measure since the company can induce trial by heavy promotion and sampling activities. In this case, the more appropriate measure is the percentage of consumers who *bought* the brand at least once. Occasionally, trial is measured by volume of sales (not percentage of consumers). Yet, this is a less desirable measure since the major objective of the trial measure is to assess the brand's penetration among the target segment. Occasionally, one might focus on not only the level of trial but also the *timing* of it, that is, slow versus fast growth. Often this is a function of the frequency of purchase in the category and the total promotional activities (of the firm and its competitors). Still, it is desirable to model it explicitly since it has major implications for the nature and schedule of the marketing mix associated with the new-product introduction. Timing of purchase is also an important consideration in defining the respondent task in concept-testing types of studies. There is considerable evidence to suggest that the longer the time horizon considered by the respondents, the higher the stated likelihood of purchasing the product. Armstrong and Overton (1971), for example, found that on the average (across different price ranges for a given product—a minicar) the intentions to purchase the product within the next four years were 52 percent higher than those for purchasing it within the next twelve months.

Repeat

The percentage of triers who bought the brand at least twice is the repeat variable. Repeat buyers can include those who buy it on a regular basis, and even among the latter group one can distinguish between those who satisfy most of their product-class requirements with the given brand versus those who satisfy only a small portion of their total requirements with the given brand. Given this heterogeneity of the repeat phenomenon, one can often distinguish between the repeat as a measure of the percentage of triers who bought the product at least once more and long-run (equilibrium-level) use of the brand. Intentions to buy are commonly used as dependent variables in concept-testing studies and predictive models based on such data. It is essential, however, to validate the results of such models and to establish the relationship between intentions and actual purchase behavior.

Other differences among forecasting models are in the number of

dependent variables employed [whether a single variable, a number of variables, or an index (such as a factor score) of a number of relevant variables], the operational definitions of these variables, and the measurement scales utilized in defining the dependent variable(s). (For a critical discussion of the measurement aspects, see Green and Tull 1978.)

Independent Variables

Five categories of independent variables can be, and have been, employed in new-product forecasting models.

Time

The simplest and most commonly used independent variable is time. These models assume that future trends in purchases will not differ markedly from past purchases. In the new-product forecasting area, they extend the sales results of a test market to the future by using any one of a number of possible functional forms. Models in this category range from simple time-series extrapolations to exponential smoothing and Markov-chain analysis.

Marketing Variables

Price, deals, advertising, in-store promotional activities, and other marketing-policy variables can be and have been, incorporated in a number of new-product forecasting models (Charnes et at. 1966, 1968; Claycamp and Liddy 1969; Massy 1969; and Urban 1968, 1969). Other models have ignored such variables and assumed some "average" marketing strategy. An example of marketing variables incorporated in some of the new-product forecasting models is presented in table 1–2.

Customer Characteristics

The adoption/behavioral forecasting models include as independent variables a number of customer characteristics. These characteristics are both situation-specific (such as attitudes toward or awareness of the given brand, product-category interest, occasion of use, and so on) and general customer characteristics (such as age, income, personality, life style, and the like).

Table 1–2
Marketing Variables Incorporated in a Few Illustrative New-Product Forecasting Models

	The Independent Variables of:			
Dependent Variable	*Demon (Charnes et al. 1966, 1968)*	*News (BBDO 1971)*	*N. W. Ayer (Claycamp & Liddy 1969)*	*Leo Burnett Early Test Market Forecasting Model (Leo Burnett, Unpublished)*
Awareness of or knowledge about product	Advertising dollars Gross impressions Reach and frequency	Product concept Creative execution Media weight Promotion	Product positioning Media impressions Copy execution Consumer promotion recall Category interest	Advertising weight (GRP) Past awareness level
Initial purchase	Advertising awareness Promotion Distribution	Promotion Distribution	Distribution Packaging Family Brand Consumer promotion Product satisfaction Category usage	Price Awareness Past trial levels
Repeat purchase			Relative price Product satisfaction repeat Purchase frequency	Repeat rate; trial level Decay rates; trial-use rate Repeat-use rate

Competitive Activities

Only a few new-product forecasting models explicitly take into account the nature and magnitude of the competitors' likely activities and retaliation against the company's new entry. Yet, to the extent that such variables are included, as in a number of unpublished forecasting simulations, they have improved the predictive accuracy of the forecast. Information on likely competitive activities is often based on management's subjective judgment. Role playing [for a review of role playing in forecasting, see Armstrong (1978, pp. 116–120)] and analysis of competitors' past behavior can be used. In addition, conjoint-analysis-based concept and product test often incorporate likely competitive actions as factors and levels in the study,

allowing for a simulation of the effect of likely competitive action. Similarly, some pretest market simulators can include simulated competitive activities.

External Environmental Forces

A few of the new-product forecasting models explicitly include environmental forces such as economic conditions, cultural and social factors, or political and legal considerations. These variables are usually *assumed* under the model's stability over space and time. They are expected to be most important for any long-run forecasting effort or when market conditions might change between the testing and the product introduction time. In these cases, the forecast should be undertaken under an explicit set of alternative future scenarios.

As with the dependent variables, forecasting models can and do differ with respect to not only the type of independent variables employed but also the number of variables, their measurement (with and without time lags), and scaling properties.

Required Data

New-product forecasting models differ considerably with respect to data requirements. A major distinction can be made based on the source of the data used, whether it consists of what people say (for example, a study of buyer intentions or expert judgment), what people do (for example, actual purchase in a test market or a pretest-market simulation), or what people have done (for example, analysis of historical data). More specifically, models differ with respect to their reliance on four types of data:

1. Most new-product forecasting models are based on primary research data. A few models are based on existing syndicated data such as MRCA or Nielsen, and a very few rely on historical data on similar products which serve as a basis for analysis by analogy.
2. Consumer response to an actual product versus response to a concept description. Whereas most forecasting models utilize data derived from consumers' response to an actual product, few are based on consumers' response to concept descriptions. These models are further removed from the "natural" market environment, but at the same time have the advantage of being undertaken prior to actual product development and at a considerably lower cost.
3. Test-market data versus pretest-market data such as data derived from

in-home test use or other consumers' surveys based on actual experience with the product. The major distinctions between these two sources of data are the "realism" of the marketing stimuli, the length of time prior to introduction required for developing the forecast, the cost involved, and degree of control over competitive and trade activities.

4. Sales response as measured, for example, by store sales (for example, Greene 1974) as opposed to consumer-response data and especially purchase and simulated purchase (for example, Blattberg and Golanty 1974).

Other important characteristics of the type of data used in new-product forecasting models are:

Panel data (for example, Ahl 1970 and Barclay 1963) versus nonpanel data (for example, Bass 1969 or Blattberg and Golanty 1978)

The specific data-collection procedure—telephone, mail, or personal interviews

The number of observations required for parameter estimation (The models range from one observation per respondent in a typical consumer survey to one to twelve months of sales data, or panel data on three or more observations per respondent.)

The cost and time required for data collection

The representativeness of the sample

Analytical Procedures

Three major aspects of analytical procedures should be considered. First, should you use a simulation or an analytical approach? Most forecasting models are based on analytical procedures such as time-series analysis and a variety of (linear and nonlinear) multiple-regression analyses. Some forecasting models involve simulations. These simulations can be used as the sole procedure or in combination with some other analytical approach. For example, some of the recent approaches for measuring consumers' evaluation of new products and services (via a conjoint-analysis procedure) were coupled with a computer simulation leading to estimates of brand shares and brand-switching patterns for the new products and concepts versus the established products.

Second, should the model parameters be estimated within a given product class (Ahl 1970) or across a number of product classes (Claycamp and Liddy 1969)? And finally, should a new-product forecasting model be

based on a single-step analysis (Herniter 1973) or on a multiple-step approach? A typical example of the latter type is the N.W. Ayer forecasting model (Claycamp and Liddy 1969) which is based on three sets of submodels. A product-knowledge submodel is the input to a trial model which, in turn, is the input to a repeat-purchase model.

Criteria for Evaluating New-Product Forecasting Models

The importance of selecting an appropriate new-product forecasting model is hardly debatable. The question, of course, is: What is meant by "appropriate"? Since there is no consensus as to the optimal characteristics of appropriate models, and since new-product forecasting models do differ on a variety of characteristics, it is essential to specify, based on the firm's specific information needs, the characteristics that the selected new-product forecasting models should have.

The obvious and most important criterion for evaluation of new-product forecasting models is the model's actual performance, that is, the accuracy of its forecast. Yet, often this criterion is of limited practical value in evaluating various new-product forecasting models since most models are claimed (by their developers) to provide an extremely high level of accuracy. A recent survey of 151 companies conducted by the National Conference Board indicated, for example, that the median forecast error for individual product line (not restricted, however, to new products) was around 5 percent. Assuming that two models can predict equally well, the question is: What other criteria can management use in evaluating alternative forecasting models? Table 1-3 presents proposed criteria for evaluating new-product forecasting models, which include (in addition to the predictive accuracy of the models) the ability to develop and implement the model and its diagnostic power.

Predictive Accuracy

To determine the predictive accuracy of a model, the user should examine not only the extent to which the forecast deviates from the actual product performance but also

1. The short- and long-run forecasting accuracy with respect to the length of time and amount of effort required to *start* the growth stage; the rate of market growth in terms of triers, repeat buyers, and volume; the level of penetration and volume when the steady-state stage is reached

Table 1–3
Criteria for Evaluating New-Product Forecasting Models

Predictive Accuracy
Short and long term
Ability to identify turning points
Economic versus statistical evaluation
Ability to Develop and Implement the Model
Technical (mathematical, statistical, and programming) skills required
Management acceptance
Data required—type, timing, and cost
Time required for model development, implementation, and maintenance
Cost of model development, implementation, and maintenance
Diagnostic Power
Forecast for defined market segment(s)
Forecast under defined marketing efforts
Forecast under defined competitive and environmental conditions
Assessment of uncertainty

2. The ability to identify turning points (reversals of trends)
3. The economic consequences of a forecasting error

The major difficulty with these considerations is, however, the frequent lack of available information on the model's long-run forecasting accuracy and its performance in identifying turning points. Similarly, information on the cost of forecasting error is not always available. If it is available, one can build into the model an estimation procedure that minimizes the cost of error.

A useful procedure for an initial validation of new-product forecasting models is to test the model against case histories on past new-product introductions. Although not a true validation, it does offer an early indication of the model's predictive ability. Such validation should be supplemented, however, with true validation efforts (on a set of data not used in estimating the model parameters).

Ability to Develop, Implement, and Maintain the Model

In evaluating a company's ability to develop, implement, and maintain a forecasting model, concern must be given to the following factors:

1. Technical skills required for the development and implementation of the model and in particular the mathematical, statistical, and programming skills.

2. Management acceptance of (including ability and willingness to use) the model. This factor is frequently associated with the *simplicity* of the model and its commonsense appeal.
3. The data required and ability to generate them within acceptable time and cost constraints.
4. The time and cost required for model development, implementation, and maintenance.

Diagnostic Power

In addition to these considerations, the diagnostic power of the model should be taken into account. Many new-product forecasting models do not provide diagnostic insights into the relative impact of various marketing strategies on the anticipated product performance. Yet, in order to be of practical value, a forecasting model should provide management with *conditional forecast* estimates, that is, anticipated forecasts for each target-market segment under various marketing strategies and alternative competitive and environmental conditions.

Recognition of the importance of the diagnostic power of new-product forecasting models is increasing and evident, for example, in a number of attempts to modify existing models to incorporate in them a conditional forecast feature (Eskin 1974; Nakanishi 1973) or take into account alternative competitive actions (Kotler 1971, chap. 4). Figure 1-2 illustrates two types of conditional forecasts—forecasts conditional on different marketing strategies (top panel) and forecasts conditional on different assumptions concerning critical exogenous variables such as market growth, competitive reactions, and government interventions, (bottom panel). Some of these effects are illustrated in the bottom panel of figure 1-2. In all these cases, obviously one should be concerned with the accuracy of the diagnostic power of the model.

Whereas forecasts conditional on the firm's marketing efforts are relatively straightforward and can be estimated quite accurately by conjoint analysis based simulators, pretest-market, and test-market data, the incorporation of competitive effects and effects of other participants in the marketing system (such as retailers) is much more complex.

Competitive date of entry, type, and level of marketing effort can have a major impact on the performance of a new product. The competitive effect can be and has been, treated in three major ways:

1. Separate estimation of the market demand (taking into consideration, however, that the market demand is also a function of the marketing effort of the firm and its likely competitors) and the market share which the firm is likely to obtain.

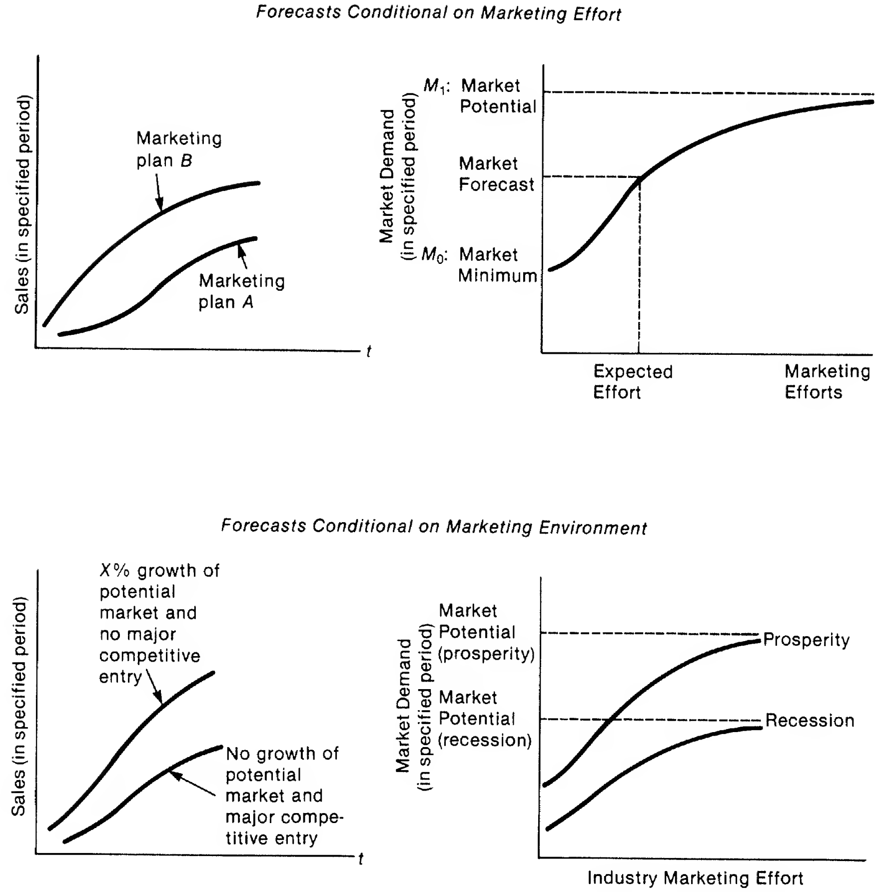


Figure 1–2. Illustrative Conditional Forecasts

2. Incorporation in a new-product forecasting model competitive-interaction parameters, such as the market-penetration coefficients of each likely competitor.
3. Simultaneous estimation of the likely performance of the firm's new product and the competitors' performance given various competitive reactions.

The competitive considerations should not be limited, however, to the forecasting of the likely competitive impact on the firm's new-product performance and should be extended to an examination of their impact on decisions such as the timing of the firm's new-product introduction and its

impact on the cost and timing of the new-product development process and the positioning of the new product and its associated marketing strategy.

Many new product forecasting models either ignore the distribution variable or include it (as in the N.W. Ayer new-product model) as one of the variables affecting trial. Yet, distribution is a critical factor affecting all facets of new product acceptance (awareness, trial, and repeat). Therefore, it is desirable to either include it in the model or develop a separate distribution (retail) forecasting model to assess the likely retail response to the firm's marketing variables—both those directed at the retailer (efforts of the sales force, trade promotions, and advertising support) and those directed at consumers (consumer advertising, price, and so on).

An important diagnostic insight that new-product forecasting models should provide is an assessment of the uncertainty involved in the new-product introduction. The specific procedures for assessing uncertainty differ according to the specific forecasting method used. For subjective judgments, for example, this can be done either by asking the respondents (experts or consumers) to assess the uncertainty of their forecast (either directly or by providing estimates for low, high, and most likely conditions) or by comparing the results of different forecasting methods. For econometric-type models, regression analysis yields useful and inexpensive estimates of uncertainty. Standard errors assess the effects of error in estimating the forecast coefficients. Cross validation on a holdout sample offers an accurate way of assessing the uncertainty of the forecast. Similarly, Monte Carlo simulations can be used to assess the effects of various types of errors.[1]

Since all forecasting methods involve a fair amount of uncertainty, the focus of forecasting developers and users should be not on a point estimate but rather on the range of possible outcomes.

It is quite rare that a single model will dominate on all these criteria. As a result, it is desirable to consider the utilization of more than one forecasting model; that is, apply the Campbell and Fiske (1959) multitrait-multimethod philosophy to the forecasting area. Following this approach would increase the user's confidence in results which appear to be consistent across two or more models. The selected models should be continuously validated and reviewed. This continuous examination may result in the modification or even replacement of existing models depending on their actual performance.

Some Illustrative New-Product Forecasting Models

Given the large number of available new-product forecasting models, no single chapter can offer a comprehensive review of all such models.[2] The

objective of this section is, therefore, to illustrate some of the major types of new-product forecasting models which vary with respect to their data bases. These include:

1. Subjective estimates by management of the new product's likely performance.
2. Analogy—finding a product with characteristics similar to those of the new product and assuming that the sales pattern of the new product will be similar to that of the analogous product.
3. Consumer-based models. These new-product forecasting models include two major sets of models—trial or first-purchase models and repeat models. These two models can be based on a number of bases, in particular concept-testing data, pre-test-market data, test-market data, and early sales data.

The following sections briefly discuss each of the three major types of models.

These three forecasting models and the various consumer-based forecasting models included in the third category vary with respect to the stage in the new-product development process for which they are most appropriate. Management's subjective estimates and analogies are usually used at the idea/concept screening stage, prior to the undertaking of a consumer-based concept-evaluation study. If this initial evaluation suggests a "go" and a concept test is conducted, the next stage is a forecasting model based on the data generated in the concept-testing study. Again, assuming a go decision, a product is developed and a simulated test market (or a simplified in-home use test) can be designed, the results of which serve as the basis for a pre-test-market forecasting model. If a go decision is reached at this stage and a test market is undertaken, the data from the test market could serve as the basis for a test-market based forecasting model. Finally, if the product is introduced to a (regional or national) market, the early-sales results can be utilized as the basis for an early-sales-based forecasting model.

The different models, despite their close association with various stages in the new-product development process, are not "pure"-type models. Past experience of marketing managers and their subjective estimates can be incorporated in other forecasting models, such as some of the concept-testing-based models. Furthermore, a forecasting model structured for a test-market type of data (requiring, for example, information on awareness, trial, repeat, and amount used by regular users) can be used at a simulated test-market stage, combining actual results of the simulated test market (on trial, repeat, and amount of use, for example) with estimates from previous introductions (analogy) on the awareness level and its dependency on the level of advertising, promotion, and distribution activities of the firm.

The models illustrated in the following sections incorporate a number of these components. They were selected, to a large extent, to encompass the key components of new-product forecasting models.

Subjective Estimates of New-Product Performance

Subjective estimates of the likely new-product performance often are used at early stages of idea/concept screening. These estimates can be based on those of the relevant marketing managers as well as of other relevant personnel and, in particular, the judgments of the sales force and other boundary personnel.

The subjective estimates vary markedly in their degree of formalism. They can range from some intuitive informal estimates suggested in a brainstorming-type session to formal and structured systems. These latter approaches include applications of the Delphi method (Linstone and Turoff 1975), the analytic hierarchy process (Saaty 1977; Wind and Saaty 1980), the decision-theoretical framework (Pengilly and Moss 1920), conditional forecasting procedures, and the design of the chain-ratio method for estimating the likely demand for a given product.

The *Delphi method* interrogates a panel of "experts" by a sequence of questionnaires. Although applied primarily to the probing of the likely occurrence of distant future events (technological developments and so on) and their impact, occasionally it has been applied to the subjective assessment of the likely performance of a new product given a variety of environmental conditions and assumptions on the marketing effort of the firm and its competitors.

The *analytic hierarchy process (AHP)* is a recent addition to the various approaches to predicting the outcome of a competitive situation (Saaty 1977). The novel aspect (and major distinction) of this approach is that it structures any complex, multiperson, multicriterion, and multiperiod problem hierarchically. By using a method for scaling the weights of the elements in each level of the hierarchy with respect to an element (for example, criterion) of the next higher level, a matrix of pairwise comparisons of the activities can be constructed in which the entries indicate the strength with which one element dominates another with respect to a given criterion.

This scaling formulation is translated into a largest-eigenvalue problem which results in a normalized and unique vector of weights for each level of the hierarchy (always with respect to the criterion in the next level), which in turn (by a principle of hierarchical composition) via a series of multiplications results in a single composite vector of weights for the entire hierarchy. This vector measures the relative priority of all entities at the lowest level that enables the accomplishment of the highest objective of the hierarchy. These relative-priority weights can provide the guidelines for the allocation

of resources among the entities at the lower levels of the hierarchy or serve as the basis for prediction.

The most notable example of the predictive power of the AHP is the prediction of the outcome of the Karpov-Korchnoi 1978 match. In this application, Saaty and Vargas (1980) combined the technical and behavioral characteristics of the two players to accurately predict the outcome of the championship game while the game was still in progress.

In the context of new-product forecasting, the AHP has been applied to the initial screening of new-product concepts (Wind and Saaty 1980). In this application, a group of managers was asked to evaluate a set of concepts on their likely performance on a set of objectives (sales, share, profitability, and so on) under a set of alternative environmental scenarios (specific levels of inflation, changes in consumer life-styles, various competitive actions, and so forth).

In addition, a series of evaluations of the interrelationships among the various concepts were undertaken to determine their interdependencies. The procedure resulted in the assignment of priorities for each of the concepts reflecting their perceived relative attractiveness to the judges with respect to the set of relevant criteria under alternative environmental scenarios.

The major attractiveness of the AHP as the conceptual and measurement approach for the early prediction of the likely performance of new-product concepts and their compatibility with the various objectives of the firm [in terms of both performance characteristics (expected share, profit, and so on) and their compatibility with the desired portfolio of products and markets)] is that it offers the following features:

A flexible formulation of the hierarchy reflecting management's value systems.

A flexible hierarchy which can incorporate *any* objectives (of varying units of measurement) and *any* courses of action (current and innovative as well as competing and complementary activities) under *any* set of environmental scenarios.

A measurement procedure based on the relevant managers' perceived relationship among the various forces, actors, actions, and personal and organizational objectives. These perceptions reflect the judges' experience and results of any previous studies and data analysis.

A built-in extension to incorporate the judgments of any number of decisionmakers and resolve conflicting views.

A flexible *process* allowing for iteration in both the structure of the problem (alternative hierarchies) and judgments.

The third approach to formal subjective forecasting of new-concepts/product performance involves the development of a *decision-theory model* which can take account of the likely consumer reactions to the new product and the likely competitive reactions. Pengilly and Moss (1920), for example, reported an early application at Proctor and Gamble in which the sales and marketing staff provided information on the likely sales of a set of new products relative to competitive products under different prices and price reactions of the competitors. These data were then submitted to a linear-programming model.

A related application of subjective judgments to the development of new-product forecasts involves a procedure similar to that used in the development of *subjective conditional forecasts.* This latter approach helps decisionmakers explicitly consider the likely effect of various variables (competitive actions, environmental conditions, and so on) on the performance of the new product. As in the case of the application of the AHP, it can lead to an open discussion of the assumptions and information held by the various participants.

A somewhat different approach to subjective forecasting is the construction of the *chain-ratio method.* This procedure involves the design of a logical sequence of likely events that culminate in a forecast of the desired behavior. Consider, for example, the following hypothetical chain ratio for the demand for a new diet-beer concept: (Population) × (personal discretionary income per capita) × (average percentage of discretionary income spent on beverages) × (average percentage of amount spent on beverages that is spent on beer) × (expected percentage of amount spent on beer that will be spent on diet beer) × (likely market share the new concept is likely to achieve). Examining this hypothetical example suggests that (1) the desired result (likely amount to be spent by consumers on the new diet-beer concept) can be achieved by a variety of chains (following somewhat different sequences of links) and (2) many of the links can be estimated, based on "hard data". Hence, this approach is quite helpful for integrating subjective judgments with hard data. Given the arbitrary nature of the suquence of links, it is desirable to construct more than one chain and compare the results of the two chains.

Analogies

Whether consumer-based information on likely new-product acceptance is incorporated in a new-product forecasting model or not, it is advisable to check the likely performance of the new product against the historical performance of similar products. If no new product in a given product class has ever achieved more than 40 percent trial and more than half of the triers

have ever continued buying it (repeat), it is unlikely that the new product under consideration will outperform this historical trend. Hence, the use of analogies can serve as a useful additional check on the results of any consumer-based new-product forecasting model.

When time or money constraints preclude the use of consumer-based studies to assess the new-product performance, one can use analogies as an approximation to the likely performance. Analogies require a careful identification of the similarities between the new product(s) and the products to serve as the base for comparison. Once the comparative products are identified, analogies can be conducted to answer two key questions: What is the likely performance of the new product? and What marketing efforts are required to achieve certain levels of performance?

To answer these questions, two different data sets are required. To assess the range of likely product performance, all that is needed is the historical performance of similar products. In this context, the narrower the definition of similarity, the better. In pharmaceutical products, for example, an analogy to new drugs is not as helpful as an analogy to drugs in the same therapeutic class and, even more specifically, to drugs in the same therapeutic class by a given market segment (specialty) and similar market conditions (level of satisfaction with existing drugs and so forth).

Figure 1–3 illustrates a series of models used as a basis for estimating the likely performance of a new product. These models should be empirically derived, but often they can be related to conceptual models. The trial curve presented in figure 1–3 is the conventional S curve. Other cases can obviously be identified, and it is not uncommon to find an exponential curve, especially when sampling and heavy promotional activities accompany the new-product introduction. Furthermore, the trial curve can be established for either households or volume. The Fourt-Woodlock (1960) forecasting model, for example, focuses on number of triers attracted, percentage of repeats, and depth of repeat. Parfitt and Collins (1968), on the other hand, rely primarily on volume trial—percentage of product-category purchasing accounted for by the new product. The volume trial automatically takes into consideration the heavy users. Since heavy users are, probabilistically, more likely to be the early triers, the volume trial curve rises more rapidly in the first few months after introduction than the household trial curve.[3] Furthermore, to a large extent it also removes the seasonality factor, since the volume trial is based on category volume, which reflects the seasonal trend among both triers and nontriers. Diagnostically, one would therefore prefer to have information on both functions.

The repeat function often takes the form of the gamma Poisson formulation which Ehrenberg (1972) termed the *NBD* (*n*egative *b*inomial *d*istribution).[4] As with trial functions, the repeat can be in terms of percentage of triers who repeat their purchases (as suggested, for example, in Fourt and

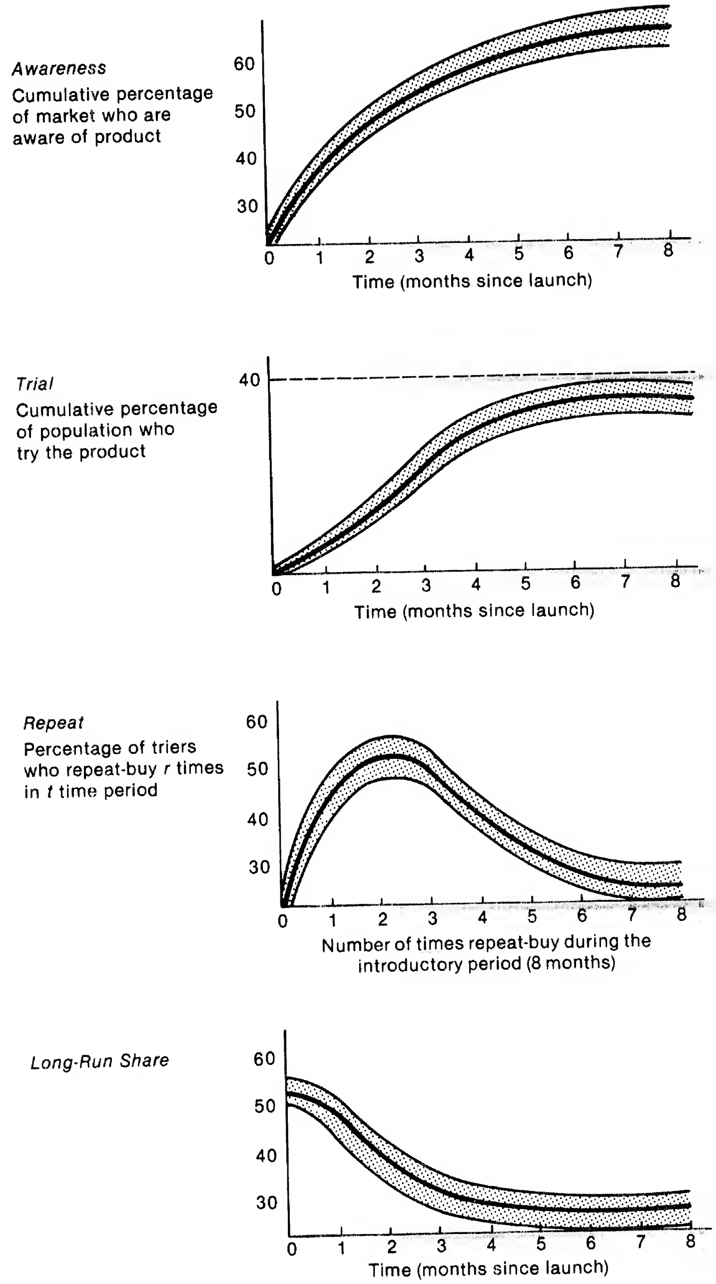


Figure 1-3. An Illustrative New-Product Performance Model Used as a Basis for Forecasting by Analogy

Woodlock 1960) in terms of volume. In this latter case (which is illustrated at the botton panel of figure 1-3) the focus is on the repeat-purchase rate—subsequent category purchases devoted to the new product—among triers who bought at least once more. The repeat-purchase rate is a key component of the Parfitt-Collins model and at the core of most repeat-purchase models.

Whereas functions such as the ones illustrated in figure 1-3 can give an idea of the likely range of performance for any new product with characteristics similar to those of the products used to develop the specific trial and repeat functions, none of these functions offer the necessary diagnostic information to assess the marketing efforts required to achieve desired performance levels. Here again, analysis by analogy can be quite helpful. Consider for example, the illustrative functions in figure 1-4. Having such empirically based functions for the various dependent variables can be of considerable diagnostic value. Ideally, one would like to have more detailed functions which are not limited to overall marketing expenditures, but rather can separate the response functions for advertising, promotion, and distribution expenditures. Data availability is a serious constraint, however, on the ability to develop such functions. Unless a company is heavily involved in new-product introduction, it can be helpful to include in the response functions the experience of one's competitors. Whereas such data can be generated, their reliability often decreases as the detailed level of analysis increases. In this context, a more serious limitation involves the control for (having accurate information on) interactions among the various marketing-mix variables, the competitive reactions to the new-product introduction, and the likely changes in relevant environmental forces.

The data requirement for forecasting by analogies can be substantial and should be reflected in the design of the firm's marketing-information system. The development of such models is a critical component of the new-product evaluation system and should be attempted even in those cases in which the firm relies heavily on primary research on the likely performance of each new product. The major benefits of such a system are that it provides an early indication of likely performance of new concepts even prior to a formal concept testing and a range of likely performance against which the results of other new-product forecasting models can be evaluated.

Consumer-Based New-Product Trial (First-Purchase) Forecasting Models

Concept-testing studies can offer the basis for excellent trial projections. Firms which have cumulative experience with concept-testing procedures

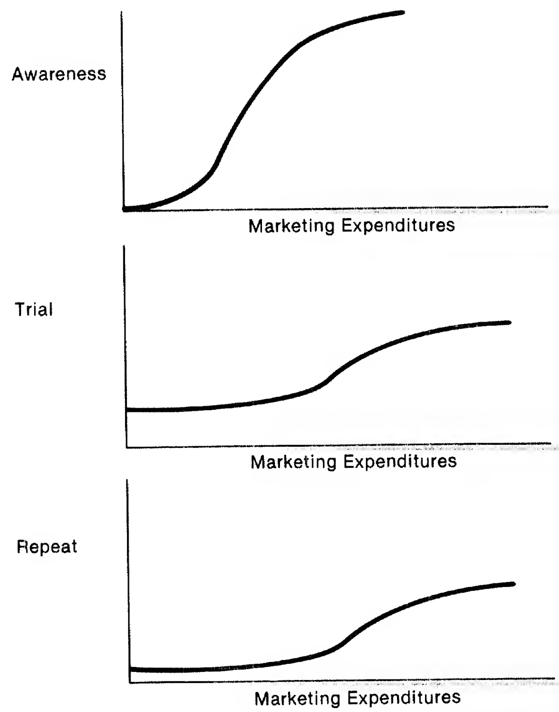


Figure 1-4. An Illustrative New-Product Market-Response Function Used as a Basis for Forecasting by Analogy

often can develop rules for translating concept-testing results to percentage of trial (in test-market conditions). Figure 1-5 illustrates such a translation rule for a frequently purchased household-product category.

Trial can be estimated quite accurately. The N.W. Ayer trial model, discussed later in this section and reproduced in chapter 14 of this book, has been reported to have high predictive accuracy. Similarly, Eskin, in an unpublished paper, reported an extremely high trial predictive accuracy using three variables: number of people who tried the product class in the last year ($\beta = .92$); availability of the brand, that is, effective distribution ($\beta = .83$); and the degree of marketing support ($\beta = .58$). The R^2 of this model was .95 (with a standard error of .15). The R^2 can be further improved by adding concept-testing scores to the model.

If a concept-testing study is based on a conjoint-analysis design, the associated computer simulation can result in estimates of trial, source of trial, and likely cannibalization. The experience with conjoint-analysis-based simulators has been quite favorable, especially when the simulation is

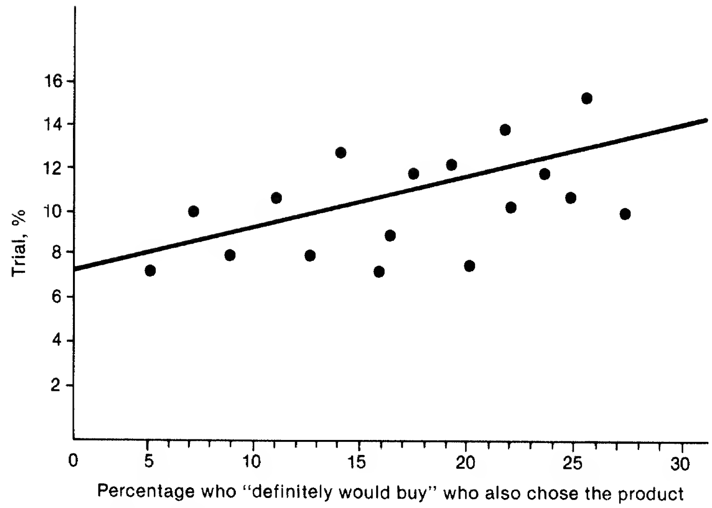


Figure 1-5. A Simplified Conversion of Concept-Testing Results to Product Trial

incorporated with a trial forecasting model (similar to those used in pre-test-market forecasting models). This includes not only information on consumer choice between various new concepts and the current products, but also information on likely level of distribution, advertising and promotion of the given brand and its competitors, and the degree of product-class penetration.

The main value of the conjoint-analysis-based trial simulators, which are also included in some of the pre-test-market models, is their ability to estimate the market share of each new concept (product) formulation (and brand) under any number of conditions, such as introduction of a new concept with no competitive retaliation; introduction of a new concept with competitive retaliation in terms of changes in their current marketing strategy or by the introduction of new or modified brands; and changes in consumer preferences for various product attributes (such as increased price sensitivity as a result of increased inflationary pressures).

Conjoint-analysis-based simulators have been employed in forecasting the likely response to new-product concepts in a variety of product categories ranging from frequently purchased consumer products for which the forecast is primarily of trial (foods, cosmetics, and so on) through consumer and industrial services (airline services, financial services, and so forth) to the design of durable consumer and industrial products (cars, office reproduction equipment, and so on).[5]

Despite the recent growth in the utilization of these models, at both the

concept-testing and the simulated test-market stage, most of the first-purchase (trial) forecasting models have been based on test-market or early-sales data.

The underlying behavioral concept of most of the trial (diffusion-type) models is the adoption-imitation process; that is, the new brand is first adopted by the "innovators," who in turn influence others to adopt it via word of mouth and demonstrated brand use.

The basic Bass (1969) type of diffusion model can be presented as

$$\frac{dN(t)}{dt} = a[(\bar{N} - N(t)] + bN(t)[\bar{N} - N(t)]$$

where $\frac{dN(t)}{dt}$ = rate of diffusion at time t [slope of $N(t)$]

$N(t)$ = cumulative number of adopters at time t

$\bar{N}$ = population of potential adopters (ceiling on number of adopters)

a = constant, coefficient of innovation (defined in table 1–4 as coefficient of external influence reflecting the influence of marketing strategy of changing agent)

b = constant, coefficient of imitation (defined in table 1–4 as coefficient of internal influence reflecting word-of-mouth effects)

The various diffusion models vary in terms of their assumptions concerning the value of a and b, whether $\bar{N}$, the market potential, is constant or not, the degree to which the diffusion model is only a function of time or also a function of the marketing strategies of the firm, the relationships of the new brand with other products in the marketplace, and whether the model takes into consideration the spatial (geographic) diffusion pattern. Table 1–4 classifies the major first-purchase diffusion models based on these variables.

Fourt and Woodlock (1960), for example, assumed that $b = 0$; that is, the diffusion is a pure innovative-effect model. Mansfield (1961) and Fisher and Pry (1971), on the other hand, assume that $a = 0$; that is, their diffusion models are pure imitation-effect models. In contrast, Bass (1969) assumes positive values for both a and b and hence incorporates both innovativeness and imitation effects.

The diffusion models vary also with respect to the way they treat the market potential. Most models (such as those of Bass, Fourt and Wood-

Table 1–4
First-Purchase Diffusion Models of New-Product Acceptance

Work by:	*Coefficient of Internal Influence*	*Coefficient of External Influence*	*Eventual Number of Buyers*
Basic Models			
Bass (1969)	Constant	Constant	Constant
Fourt and Woodlock (1960)	0	Constant	Constant
Mansfield (1961)	Constant	0	Constant
Hendry (1972)	Constant	0	Constant
Lekvall and Wahlbin (1973)	Constant	Constant	Constant
Lawrence and Lawton (1980)	Constant	Constant	Constant
Extensions			
Robinson and Lakhani (1975)	*f* (price)	Constant	Constant
Horsky and Simon (1978)	Constant	*f* (advertising)	Constant
Hernes (1976)	*f* (time)	*f* (time)	Constant
Bass (1978)	*f* (demand elasticity, learning parameters, price)	*f* (demand elasticity, learning parameters, price)	Constant
Lilien and Rao (1978)	Constant	*f* (personal selling)	Constant
Peterson and Mahajan (1978)	*f* (product relationships)	Constant	Constant
Mahajan and Peterson (1978)	Constant	Constant	*f* (all relevant variables)
Mahajan et al. (1979)	Constant	0	*f* (housing starts)
Dodson and Muller (1978)	0	Constant	*f* (advertising)
Chow (1967)	Constant	0	*f* (price)
Lackman (1978)	Constant	0	*f* (profit/sales)

Source: V. Mahajan and E. Muller, "Innovation Diffusion and New Product Growth Models," *Journal of Marketing.* 43 (Fall 1979):55–68. Diffusion model; $n(t) = a\,\Delta N(t) + bN(t)\,\Delta N(t)$ where a is coefficient of external influence, b is coefficient of internal influence, $\Delta N(t)$ is the difference between eventual number of adopters and current number of adopters. And $\Delta N(t) = \bar{N} - N(t)$ for all the above except for Hendry (1972), Chow (1967), and Lackman (1978), where $\Delta N(t) = \ln\bar{N} - \ln N(t)$. Also $N(t)$ is the cumulative number of adopters at time t, and $\bar{N}$ is the eventual number of adopters.

lock, and Fisher and Pry) assume that it is constant over time (either as the total population or a fraction of it), while others, such as Mahajan and Peterson (1978) and Dodson and Muller (1978), assume that the potential adopter population changes over time, that is, $\bar{N}(t) = f(S(t))$, where $S(t)$ is a vector of all relevant exogenous and endogenous variables affecting $\bar{N}(t)$.

The pioneering diffusion models of Bass (1969) and Fourt and Woodlock (1960) do not include marketing-strategy effects. This omission has been overcome in a number of extensions of the basic diffusion model. Robinson and Lakhani (1975) incorporate marketing variables by representing *b*, the imitation coefficient, as a function of marketing-strategy variables such as advertising, promotion, and price. More specifically, they represent *b* in the basic diffusion model as a function of price. Horsky and Simon (1978) incorporated the marketing-strategy variables not through their effect on *b*, but on innovators—*a*. They assume advertising serves as a source of information to innovators and hence affects the coefficient of innovation *a*.

Further extension of the basic diffusion model was suggested by Peterson and Mahajan (1978), who incorporated in the model the complementarity, substitutability, contingent, and independent relations of the new brand with other brands in the marketplace. The substitution effect, for example, is incorporated in the basic diffusion model by adding the constant *c*, which represents the substitution effect—the interaction between the adopters of product *j* and the nonadopters of product *i*—which results in a decrease in the rate of diffusion for product *i*. This effect can be presented as

$$\frac{dN_i(t)}{dt} = a_i[\bar{N}_i - N_i(t)] + b_i N_i(t)[\bar{N}_i - N_i(t)] - c_i N_j(t)[\bar{N}_i - N_i(t)]$$

Most of the diffusion models ignored the spatial diffusion pattern. Yet, given the pattern of new-product introduction which often follows the regional pattern (even when national introduction is planned, not all areas receive the products at the same time and with the same intensity), incorporating a spatial diffusion pattern is most desirable. Mahajan and Peterson (1979), for example, offer a limited spatial extension of the basic diffusion model. In the application, they expressed *a, b,* and *N* as a function of distance, but limited the spatial diffusion pattern to an initial product introduction to one region.

In addition to the basic diffusion model and its various extensions and the conjoint analysis based first-purchase forecasts, the forecast of a new-product trial has also been the subject of extensive econometric modeling efforts. One of the most widely used models in this category for frequently purchased products is the N.W. Ayer model (Claycamp and Liddy 1969):

$$IP = a_2 + b_{21}(AR) + b_{22}(DN*PK) + b_{23}(FB) + b_{24}(CP) + b_{25}(PS*)$$
$$+ b_{26}(CU) + e$$

where IP = initial purchase

AR = advertising recall, which is a function of product positioning (PP), media impression (AHI), copy execution (CE), consumer promotion (CP), and category intent (CI) and takes the form of

$$AR = a_1 + b_{11}(PP) + b_{12}\sqrt{AHI \times CE} + b_{13}(CP) + b_{14}(CI) + e$$

DN = distribution

PK = packaging

FB = family brand

CP = consumer promotion

PS = product satisfaction

CU = category use

e = error

The specific regression equation identified for this model, based on over 100 cases, is

Initial purchase = -16 + 0.19(distribution $\times$ packaging) + 9.25(family brand) + 0.09(consumer promotion) + 0.02(product satisfaction) +0.07(category use) + 0.37(predicted advertising recall)

and the function for advertising recall is in turn:

Recall = -36 + 0.76(product positioning)
+ 2.12 $\sqrt{\text{media impressions} \times \text{copy execution}}$
+ 0.04(consumer promotion recall) + 0.39(category interest)

Consumer-Based Repeat-Purchase and Share Forecasting Models

When the long-run performance of a product depends on repurchase (including cases such as the need for renewal of contract), a first-purchase

(trial) forecasting model should be supplemented with a repeat-purchase model. The separation of sales into trial and repeat is critical, given that the same level of sales x can be reached by various combinations of trial and repeat, including the extreme disastrous cases of only trial and no repeat. Furthermore, having information on the trial/repeat breakdown offers critical diagnostic insights. Consider, for example, the case of very low trial and high repeat, which suggests the need to change the introductory marketing strategy to generate higher trial levels.

Some of the various trial/repeat combinations for a given level of sales are illustrated in the following table:

Total sales in period t	=	Number of first-time buyers in period t	×	Average purchase volume per period per time t	+	Number of repeat buyers in period t	×	Average purchase volume per period per repeat customer
100	=	100		1		0		0
100	=	20		1		10		8
100	=	50		1		25		2
100	=	70		1		30		1

An examination of these four cases suggests that it is helpful to focus on the number of triers and repeaters, and to the extent that there are any differences in the average purchase volume of the two (as is often the case), these data should also be included.

The focus on repeat should not be limited, however, to the number (or proportion) of first purchasers. Fourt and Woodlock (1960), Massy (1969), Eskin (1973), and most other repeat-purchase models recognize the need to model the *depth of repeat* (the repeat purchase of one-time buyers, two-time buyers, and so on). The various repeat-purchase models vary, however, in their specific formulation. Table 1–5 briefly classifies most of the published repeat-purchase models on a number of critical characteristics. Three of these models are briefly discussed.

First, the Fourt and Woodlock (1960) is based on the concept of repeat ratio. The first repeat ratio, for example, is the fraction of triers who make a second purchase. Second, third, and other repeat ratios are similarly defined and can be interpreted as the probability that a first- or second- or third-time buyer will buy again. The repeat-sales equation used in this model is

$$S_t = S_{F_t} N_{F_t} \left[1 + \frac{N_{1_t}}{N_{F_t}} + \frac{N_{1_t}}{N_{F_t}} \frac{N_{2_t}}{N_{1_t}} + \frac{N_{1_t}}{N_{F_t}} \frac{N_{2_t}}{N_{1_t}} \frac{N_{3_t}}{N_{2_t}} + \cdots \right]$$

where S_t = total sales in period t

S_{F_t} = average purchase volume per period per first-time buyer

N_{F_t} = number of first-time buyer

$\frac{N_{1_t}}{N_{F_t}}$ = first repeat ratio (fraction of triers who made a second purchase)

$\frac{N_{2_t}}{N_{1_t}}$ = second repeat ratio (fraction of first repeaters who made a second repeat purchase)

This model is often used by the National Purchase Diary Panel, Inc. (NPD) to predict first-year volume. Their experience with it is not satisfactory, however, in predicting the long-run sustaining volume, for which they use the Parfitt and Collins model.

Second, the Parfitt and Collins (1968) model, as the Fourt and Woodlock model, requires panel data. It does differ, however, in its formulation. It is not based on the number of new buyers each period. Rather, the Parfitt and Collins model is designed for long-run share (S) predictions as a function of three factors:

$$S = prb$$

where p = ultimate penetration rate of brand (percentage of product-class buyers who try brand)

r = ultimate repeat-purchase rate (new-brand purchases as percentage of all purchases by persons who once purchased this brand)

b = buying-rate index of repeat purchase of this brand (when average buyer = 1.00)

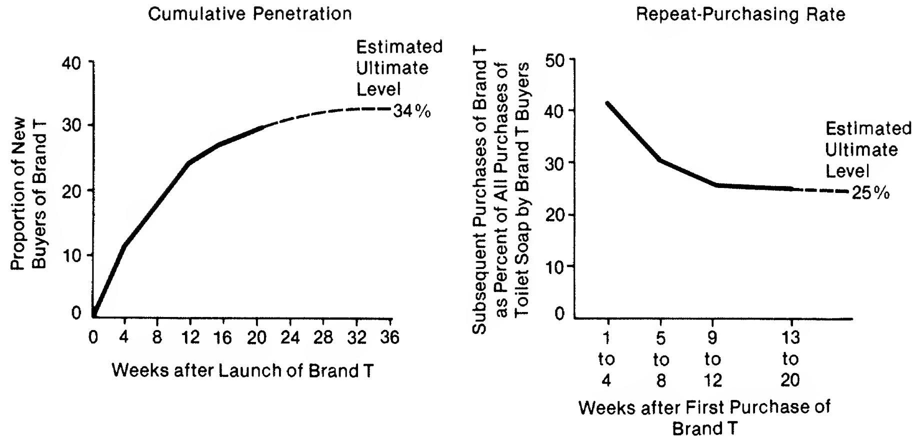
By using this formulation, if, for example, 34 percent of buyers in this market ultimately try the brand (p = 0.34) and 25 percent of their subsequent repurchases go to this brand (r = 0.25) and those buying the brand buy an average quantity (b = 1.00), then the predicted long-run equilibrium brand share would be 8.5 percent [(0.34)(0.25)(1.00)]. If the brand attracts on average heavier buyers (b = 1.20), then the predicted share would be higher—10.2 [(0.34)(0.25)(1.20)].

These ultimate share predictions can be made as soon as the penetration curve and repeat-purchase curve stabilize (tend to move toward an asymptotic value). This is illustrated in figure 1-6 and avoids the need to wait until the stabilization of share.

Table 1–5
Repeat-Purchase Models

			Distinguishes Between			*Includes*			
Model	*Deter-ministic/ Stochastic*	*Integrated Process*	*Unaware Potential*	*Types of Infor-mation*	*Depth of Repeat*	*Word of Mouth*	*Impulse Purchase*	*Decay*	*Recycle*
Fourt and Woodlock (1960)	D				✓				
Parfitt and Collins (1968)	D								
Claycamp and Liddy (1969): Ayer	D		✓						
Nakanishi (1973)	S	✓			✓			✓	
Massy (1969): STEAM	S	✓			✓				
Assmus (1975) NEWPROD	D		✓					Partial	Partial
Urban (1970) SPRINTER	D	✓	✓		Partial	✓		Partial	✓
Midgley (1976)	D	✓		✓		✓			✓
Blattberg and Golanty (1978): TRACKER	D		✓		✓			Partial	
Dodson and Muller (1978)	D	✓	✓			✓		✓	✓
Lilien and Rao (1978)	D	✓				✓		✓	

Source: V. Mahajan and E. Muller, "Innovative Behavior and Repeat Purchase Diffusion Models" (Wharton School Working Paper, Philadelphia, Pa. 1980).



Source: J.H. Parfitt and B.J.K. Collins, "Use of Consumer Panels for Brand-Share Predictions," *Journal of Marketing Research* 5(May 1968): 132, 133.

Figure 1–6. Illustrative Cumulative Penetration and Repeat-Purchasing Rate

The Fourt and Woodlock and the Parfitt and Collins models do not include explicitly any marketing-strategy variables. Parfitt and Collins, for example, stated explicitly the two key assumptions of their model:

1. The retail distribution of the new brand is uniformly high in the area under study; or, failing that, it is not substantially worse now than it is likely to be in the foreseeable future.
2. Besides the advertising and promotional activity accompanying and during the brand's launch (including competitor's retaliatory measures), the circumstances of the market will remain much the same in the future as they have been during the prediction measurement.

Third, a model which develops a marketing-strategy-dependent repeat-buying model is the N.W. Ayer model. According to this model, which has not been successfully validated, however, the repeat purchase is a function of four variables: the initial purchase prediction (which, as discussed earlier, incorporated all the marketing variables except price), relative price, product satisfaction, and purchase frequency.

Most repeat-purchase models use the concept of depth of repeat. These models differ, however, in their formulation and complexity [from the relatively simple Fourt and Woodlock model through the sophisticated and somewhat more complex Eskin (1973) model and most complex STEAM model of Massy (1969), which presumes population heterogeneity and introduces the concept of depth of trial classes and different propen-

sities to enter them]. In addition, repeat forecasting models vary in the way they are incorporated in long-run share and use forecasts. As with first-purchase (trial) models, most models are claimed (by their developers) to be highly valid. The user of these models should, therefore, select the one which best fits his or her specific need or, alternatively, build a model using some of the more desirable components from the currently available models.

Incorporating a Forecasting Model in an Economic-Evaluation System

The output of new-product forecasting procedures is rarely considered in isolation from other factors involved in the go/no go decision. It is important, therefore, to consider a formal integration of the new product forecasting methods with the firm's economic-evaluation system. One of the more comprehensive systems that incorporates forecasting and economic evaluation is the DEMON (*de*cision *m*apping via *o*ptimum go/no go *n*etworks) model developed to aid management in evaluating alternative plans for the introduction of new products and alternative marketing-research studies that could help improve the product's profitability. The marketing planning framework of DEMON and its key functions are presented in figure 1-1a. These functions, which were determined by a least-squares regression on data available on over 200 packaged-goods products, are oversimplistic since they ignore other possible determinants of the key dependent variables. Trial, for example, can be affected by advertising frequency, promotion activities, distribution, and price. These modifications can easily be incorporated, however, in the model. The unique feature of the DEMON system is its focus on the search for an optimal path through a total information network subject to management's specified constraints concerning the payback period, the planning horizon, mimimum acceptable profits for a go and on decision, minimum degree of confidence needed for a go and on decision, and total marketing-research budget.

The model evaluates the various (research) patterns to a "go" and "on" decision by calculating the estimated profit (based on the cost and demand estimates) and risk for each path. The "best" path is selected, and if an "on" decision (specific study) were recommended, the next step would be an evaluation using the results of this study.

The SPRINTER (*s*pecification of *pr*ofits with *in*teractions under *t*rial and *e*rror *r*esponse) model (Urban 1970) is similar in its objective to DEMON with the added feature of considering relevant interactions of the new product with the other products of the firm.

DEMON, SPRINTER, and the more recent POSSE system (Green and Carroll 1980) all offer formal procedures for incorporating the results of new-product forecasting models in an economic-evaluation system. Whether these specific modeling efforts or others are used is immaterial. The key to appropriate use of new-product forecasting models is, however, their incorporation in some formal economic-evaluation system that could translate trial, repeat, intentions to buy, and other results of forecasting models into go, no go, or on decisions.

Conclusions

No single best forecasting model has been identified yet. Each of the numerous new-product forecasting models is based on specific assumptions and data requirements. Since the data available and the accuracy needs of new-product forecasting models vary by the specific decision for which the forecast is needed, it is desirable to develop a system of forecasting models. Such a system (composed of a number of diverse forecasting methods) would ideally be based on at least four concepts:

1. *The inclusion of the most appropriate forecasting method for each stage of the new-product development process.* Given the cost and value of information at early stages of development, it would be desirable to have a relatively cheap and quick forecasting procedure with lower accuracy. Such procedures should offer input to a simple go/no go decision by assessing the likelihood that the new product's performance would fall within some acceptable range. In contrast, at later stages of development, much greater accuracy is needed and hence different forecasting models should be considered. This concept is not unlike the evolutionary model building proposed by Urban and Karash (1971).
2. *The inclusion of at least two forecasting methods at each stage of the new-product development process.* This multiple-method approach would offer a sort of cross validation of the outcomes and would increase management's confidence in the forecasting results.
3. *An explicit make-buy decision.* In selecting new-product forecasting models, one should not be constrained to existing commercial models. A make-buy decision should be made based on an examination of the available commercial models (such as the N.W. Ayer trial model, the ASSESSOR pretest-market simulator, and others) versus the advantages of designing an idiosyncratic forecasting model for the firm which could incorporate the most attractive features of the various commercial models.
4. *Avoid overly sophisticated models.* Many of the new-product fore-

casting models, not unlike some of the ones described in this chapter, tend to err on being overly sophisticated and mathematically elegant. Yet, an interesting finding of two recent evaluations of various forecasting methods (Armstrong 1978 and Ascher 1978) suggests that the more sophisticated model does not necessarily lead to more useful forecasts.

Forecasting the likely performance of new concepts or products is a critical step of every new-product development effort. It should be an integral part of the new-product development system, and continuous attention should be given to the modification and improvement of the forecasting system so as to ensure a reliable and valid forecasting methodology. This requires continuous followup and examination of the statistical and economic accuracy of the system.

Since *new-product* forecasting is a complex and difficult task, improvement of this methodology and its implementation could have secondary benefits to the other forecasting tasks facing a firm (environmental forecasting and so on).

Notes

1. For a discussion of these and other approaches to assessing the uncertainty of various forecasting methods, see Armstrong (1978).
2. For reviews of some of these models, see Kotler (1971, chap. 17) and Rao and Cox (1978).
3. A number of studies among physicians clearly indicate that the heavy-drug-category users are the most likely to be the first to try any new drug in the given category. This finding, as it relates to consumer populations, was also found by Parfitt and Collins (1968).
4. The NBD—gamma Poisson formulation assumes that different repeat buyers have different long-run rates of buying and that these rates follow a gamma distribution. The Poisson part assumes that in the short run each buyer will randomly deviate from his particular long-run rate so that his chance of any given purchase frequency follows the Poisson distribution.
5. For a more complete list of areas of conjoint-analysis applications, see Green and Srinivasan (1978) and Wind (1978).

References

Ahl, David A. "New Product Forecasting Using Consumer Panels." *Journal of Marketing Research* 7(May 1970): 160–167.

Armstrong, Scott J. *Long-Range Forecasting: From Crystal Ball to Computer.* New York: Wiley-Interscience, 1978.

Armstrong, Scott J. and Overton, Terry. "Brief vs. Comprehensive Description in Measuring Intention to Purchase." *Journal of Marketing Research* 8(February 1971): 114–117.

Ascher, William. *Forecasting: An Appraisal for Policy Makers and Planners.* Baltimore, Md.: Johns Hopkins University Press, 1978.

Assmus, G. "NEWPROD: The Design and Implementation of a New Product Model." *Journal of Marketing Research* 39(January 1975): 16–23.

Bailey, Norman T.J. *The Mathematical Theory of Epidemics.* New York: Hafner Publishing Co., 1957.

Barclay, William D. "A Probability Model for Early Prediction of New Product Success." *Journal of Marketing* 27(January 1963): 63–68.

Bass, Frank M. "A New Product Growth Model for Consumer Durables." *Management Science* 15(January 1969): 215–227. [ch. 19 of this book.]

———. "An Integration of the New Product Growth Model with the Experience Cost Function and Optimal Pricing." Presented at the The Institute of Management Science/Operations Research Society of America national meeting in New York City, May 1978.

BBDO, Management Science Department. *The Theoretical Basis of NEWS.* New York: BBDO, unpublished 1971.

Blattberg, Robert, and John Golanty. "An Early Test Market Forecasting Model for New Product Planning." Working Paper, University of Chicago, 1974.

———. "TRACKER: An Early Test Market Forecasting and Diagnostic Model for New Product Planning." *Journal of Marketing Research* 15(May 1978): 192–202. [ch. 16 of this book.]

Brown, D.A.; Buck, S.F.; and Pyatt, F.G. "Improving the Sales Forecast for Consumer Durables." *Journal of Marketing Research* 2(August 1965): 229–235.

Burger, Philip C. "Developing Forecasting Models for New Product Introductions." *AMA Proceedings,* Fall 1968, pp. 112–118.

———. "COMP: A New Product Forecasting System." Working Paper, Northwestern University, Evanston, Ill., 1972. [A revised version is included in ch. 10 of this book.]

Campbell, Donald T., and Fiske, D.W. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56(1959).

Charnes, A.; Cooper, W.; DeVoe, J.; and Learner, D. "DEMON: Decision Making Via Optimum Go-No Go Networks—A Model for Marketing New Products." *Management Science* 12(July 1966): 865–887.

———. "DEMON Mark II: External Equations, Solution and Approximation." *Management Science* 14(July 1968): 682–691.

Chow, G.C. "Technological Change and the Demand for Computers." *American Economic Review* 57(December 1967): 1117–1130.

Cima, F. "A Forecasting Model for the Market of Competitive Products." *Technological Forecasting and Social Change* 4(1973): 51–65.

Claycamp, H.J., and Liddy, L.E. "Prediction of New Product Performance: An Analytical Approach." *Journal of Marketing Research* 6(November 1969): 414–420. [ch. 14 of this book.]

Dodson, J.A., and Muller, E. "Models of New Product Diffusion through Advertising and Word-of-Mouth." *Management Science* 15(November 1978): 1568–1578.

Ehrenberg, A.S.C. "Predicting the Performance of New Brands." *Journal of Advertising Research* 11(November–December 1971): 3–10.

———. *Repeat Buying: Theory and Applications.* New York: American Elsevier Publishing Company, Inc., 1972.

Eskin, Gerald J. "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model. *Journal of Marketing Research* 10(May 1973): 115–129.

———. "Casual Structure in Dynamic Trial-Repeat Forecasting Models." Paper presented at the 57th International Marketing Conference, April 15–18, 1974, Montreal.

Fisher, J.C., and Pry, R.H. "A Simple Substitution Model for Technological Change." *Technological Forecasting and Social Change* 2(1971): 75–88.

Fourt, Louis A., and Woodlock, Joseph W. "Early Prediction of Market Success for New Grocery Products. *Journal of Marketing* 24(October 1960): 31–38.

Green, Paul E., and Carroll, Douglas J. "New Computer Tools for Product Strategy." University of Pennsylvania Working Paper, April 1980. [ch. 4 of this book.]

Green, Paul E., and Srinivasan, V. "Conjoint Analysis in Consumer Research: Issues and Outlook." *Journal of Consumer Research* 5(September 1978): 103–123.

Green, Paul E., and Tull, Donald S. *Research for Marketing Decisions,* 4th ed. Englewood Cliffs, N.J.: Prentice-Hall, 1978.

Greene, J.D. "Projecting Test Market 'Trial-Repeat' of a New Brand in Time." Talk to the American Marketing Association, 1974.

Griliches, Zvi. "Hybrid Corn—An Exploration in the Economics of Technical Change." *Econometrica,* October 1967, pp. 501–522.

Hamburg, Morris, and Atkins, R.J. "Computer Model for New Product Demand." *Harvard Business Review* 45(March–April 1967): 107–113.

Hendry, I. "The Three Parameter Approach to Long Run Forecasting." *Long Range Planning* 5(March 1972): 40–45.

Hernes, F., "Diffusion and Growth—The Non-Homogeneous Case." *Scandinavian Journal of Economics* 78(1976): 427–436.

Herniter, Jerome D. "An Entropy Model of Brand Purchase Behavior." *Journal of Marketing Research* 10(November 1973): 361–375.

———. "A Comparison of the Entropy Model and the Hendry Model." *Journal of Marketing Research* 11(February 1974): 21–29.

Horsky, D., and Simon, L.S. "Advertising in a Model of New Product Diffusion." Presented at the The Institute of Management Science/ Operations Research Society of America national meeting in New York City, May 1978.

Kalwani, M., and Silk, A. "Structure of Repeat Buying for New Packaged Goods." *Journal of Marketing Research* 17(August 1980): 316–322.

Kotler, P. *Marketing Decision Making: A Model Building Approach.* New York: Holt, Rinehart & Winston, 1971.

Lackman, C.L. "Gompertz Curve Forecasting: A New Product Application." *Journal of the Market Research Society* 20(January 1978): 45–47.

Lawrence, Kenneth D., and Lawton, William H. "Product Life Cycle Forecasting in Industrial Marketing." [ch. 22 of this book.]

Learner, D.B. "Profit Maximization through New Product Market Planning and Control." In *Applications of the Sciences in Marketing Management*, edited by F.M. Bass, C.W. King, and E.A. Pessemier. New York: John Wiley & Sons, 1968, pp. 151–167.

Lekvall, P., and Wahlbin, C. "A Study of Some Assumptions Underlying Innovation Diffusion Functions." *Swedish Journal of Economics* 75(1973): 362–377.

Lilien, G.L., and Rao, A.G. "A Marketing Promotion Model with Word-of-Mouth Effects." Working Paper 976–78, Sloan School of Management, Massachusetts Institute of Technology, Boston, Mass., 1978.

Linstone, H., and Turoff, M. *The Delphi Method: Technniques and Applications.* Reading, Mass.: Addison-Wesley, 1975.

Lipstein, Benjamin. "Modeling and New Product Birth." *Journal of Advertising Research* 10(October 1970): 3–11.

Mahajan, V., and Muller, E. "Innovation Diffusion and New Product Growth Models." *Journal of Marketing* 42(Fall 1979): 55–68. [ch. 18 of this book.]

———. "Innovative Behavior and Repeat Purchase Diffusion Models." Wharton School Working Paper, Philadelphia, Pa. 1980.

Mahajan, V., and Peterson, R. "Innovation Diffusion in a Dynamic Potential Adopter Population." *Management Science* 24(November 1978): 1589–1597.

———. "Integration Time and Space in Technological Substitution Models." *Technological Forecasting and Social Change* 14(August 1979): 231–241.

Mahajan, V.; Peterson, R.; Jain, A.K.; and Malhotra, N. "A New Product

Growth Model with a Dynamic Market Potential." *Long Range Planning* 12(August 1979): 51–58.

Mansfield, Edwin. "Technical Change and the Rate of Imitation." *Econometrica,* October 1961, pp. 741–766.

Massy, William F. "Forecasting the Demand for New Convenience Products." *Journal of Marketing Research* 6(November 1969): 405–413.

Midgley, D.F. "A Simple Mathematical Theory of Innovative Behavior." *Journal of Consumer Research* 3(June 1976): 31–41. [ch. 20 of this book.]

Nakanishi, Masao. "Advertising and Promotion Efforts on Consumer Response to New Products." *Journal of Marketing Research* 10 (August 1973): 242–249.

Parfitt, J.H., and Collins, B.J.K. "Use of Consumer Panels for Brand-Share Predictions." *Journal of Marketing Research* 5(May 1968): 131–145. [ch. 13 of this book.]

Pengilly, P.J., and Moss, A.J. "Choice of a New Product, Its Selling Pattern Price." *Operational Research Quarterly,* 1920, pp. 179–185.

Peterson, R.A., and Mahajan, V. "Multi-Product Growth Models." In *Research in Marketing,* edited by J. Sheth. Greenwich, Conn.: JAI Press, 1978.

Rao, V., and Cox, J.E., Jr., *Sales Forecasting Methods: A Survey of Recent Developments.* Monograph #78–119. Boston: Marketing Science Institute, 1978.

Robinson, V., and Lakhani, C. "Dynamic Price Models for New Product Planning." *Management Science* 21(June 1975): 1113–1132.

Rogers, Everett M. *Diffusion and Innovations.* New York: The Free Press, 1962.

Ryans, A.F. "Estimating Consumer Preferences for a New Durable Brand in an Established Product Class." *Journal of Marketing Research* 11(November 1974): 434–443.

Saaty, Thomas L. "A Scaling Method for Priorities in Hierarchical Structures." *Journal of Mathematical Psychology* 15(June 1977): 234–281.

Saaty, T., and Vargas, Luis G. "Hierarchical Analysis of Behavior in Competitive Prediction in Chess." *Behavioral Science* 25(May 1980): 180–191.

Urban, Glen L. "A New Product Analysis and Decision Model." *Management Science* 11(April 1968): 490–517.

———. "New Product Growth Model for Consumer Durables." *Management Science* 12(January 1969): 215–227.

———. "SPRINTER MOD III: A Model for the Analysis of New Frequently Purchased Consumer Products." *Operations Research* 18(September 1970): 805–855.

Urban, G., and Karash, R. "Evolutionary Model Building." *Journal of*

Marketing Research 8(February 1971): 62–66.
Wind, Yoram. "New Developments in Conjoint Analysis." Presented at the 25th Annual Midwest Conference of the American Statistical Association, Chicago, Ill., March 1978.
Wind, Yoram, and Saaty, Thomas. "Marketing Applications of the Analytic Hierarchy Process." *Management Science* 26(July 1980): 641–658.

Part II
Concept-Test-Based Approaches to New-Product Forecasting

The first consumer-based new-product forecast is often developed at the idea-concept stage. Idea/concept test types of studies tend to provide answers to questions such as: What is the consumer reaction to the concept? What is the size of the potential market? Which attributes are most important in accounting for consumer reactions to the concepts? More recently many concept-testing studies have supplemented this basic information with more diagnostic data from the segmentation/positioning decision, that is, they attempt to answer the questions: How many market segments exist, and what are their characteristics (socioeconomic, demographic, psychographic, and product and brand-use patterns) and reactions to the various concepts? What is the relevant competitive setting of the given concept? What is the most desirable positioning for the concept, and how likely is the concept to cannibalize the current products of the firm?

The first chapter in this part, by Shocker and Srinivasan, presents an overview of the multiattribute approaches to concept evaluation and generation. The approaches reviewed generally fit within a framework where products or brands are represented abstractly in terms of attributes potentially relevant to consumer choice. Competing brands are indicated by their perceived locations (coordinates or levels) in a "perceptual product space" whose axes are relatable to such attributes. The approaches model consumer preferences for (or choices among) these brands in terms of multiattribute models whose parameters vary for different conditions of intended use.

One of the major multiattibrute approaches to the modeling and forecasting of consumer reactions to new-concept offerings is conjoint analysis. To date, over 500 commercial studies have utilized this approach primarily for concept and product testing. One of the early expository papers, by Green and Wind, is included as chapter 3. This chapter provides a brief, nontechnical description of conjoint analysis and some of its early applications to concept testing. The appropriateness of conjoint analysis to concept testing was recognized as early as 1973 (Green and Wind 1973; Wind 1973). Most of these studies incorporate the utility functions in a computer simulation aimed at assessing the expected share of choices and brand-switching patterns for various concept configurations (Wind, Jolly, and O'Connor 1975). The simulations offer management the ability to

answer two major questions: For any given concept description (defined as a combination of attributes), which market segment is most attracted to it? and for any given market segment (defined in terms of demographic, psychographic, or brand-use characteristics), which is the "best" concept description?

The second generation of conjoint-analysis-based new-product forecasting models is briefly described in chapter 4 by Green and Carroll. In this chapter the authors outline their POSSE system which allows for simpler data collection than that commonly used in conjoint-analysis-type studies and offers, in addition to the conventional simulations, an optimization procedure and a system for sensitivity analysis of the results. Chapter 4 also includes an appendix, especially written for this book, which describes the various computer modules of the POSSE system.

Many concept-testing-based new-product forecasting models focus on a categorical response such as adoption or nonadoption. Logit-type models for modeling such behavior have been proposed (Green, Carmone, and Wachspress 1977). More recently Flath and Leonard compared the Green, Carmone, and Wachspress model to a maximum-likelihood logit-type model. Their comparison is found in chapter 5 and illustrates how such models can be used in forecasting a categorical response to a new concept.

Chapter 6, by Tauber, presents an industry perspective and suggests that the traditional approaches to concept testing—survey of purchase intentions—are still most popular in industry's application of concept-testing studies. Yet, in evaluating this assessment, the reader should recall that the more sophisticated new-concept forecasting models discussed in part II all have been utilized by a large number of consumer, industrial, and service firms (Myers, Greyser, and Massy 1979; Wind 1981).

References

Green, Paul E.; Carmone, Frank; and Wachspress, David P. "On the Analysis of Qualitative Data in Marketing Research." *Journal of Marketing Research* 14(February 1977):52–59.

Green, Paul E., and Wind, Y. *Multi-Attribute Decisions in Marketing.* Hinsdale, Ill.: The Dryden Press, 1973.

Myers, John G.; Greyser, Stephen A.; and Massy, William F. "The Effectiveness of Marketing's R&D for Marketing Management: An Assessment." *Journal of Marketing* 43(January 1979):17–29.

Wind, Yoram. "A New Procedure for Concept Evaluation." *Journal of Marketing* 37(October 1973):2–11.

———. *Product Policy: Concepts, Methods, and Strategy*. Reading, Mass.: Addison-Wesley, 1981.

Wind, Yoram; Jolly, Stuart; O'Connor, Arthur. "Concept Testing as Input to Strategic Marketing Simulations." In *Proceedings of the 58th International AMA Conference,* edited by E. Mazzie. Chicago: American Marketing Association, 1975, pp. 120–124.

2 Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review

Allan D. Shocker and
V. Srinivasan

For years a disparity has been apparent between the philosophy of marketing decision making and its practice. The textbook definition of the marketing concept demands the firm work from an understanding of the wants and desires of some reasonably well-defined target market to the design of its market offerings. When one observes the practice of many firms, however, the process appears reversed. The initial design of market plans often is accomplished without explicit customer inputs. Potential customers are asked to react to such plans only at later stages in their development (if asked at all) and such inputs are used only to guide management in selecting from a small number of proposals for implementation or in modifying an already developed plan (Day 1968; Stefflre 1971). Increased implementability of the marketing concept appears to wait upon an improved ability to identify relevant market desires and an understanding of how people translate them into purchase behavior. Such understanding must be based on recognition of the customer's natural ability to react to alternatives rather than create them, and must be communicated to the marketing planner in a manner facilitating translation into market actions.

Today such a research mode is emerging. It promises to aid implementation of the marketing concept by permitting customer inputs to be used at the earliest stages in the development and selection of a broad range of marketing strategies. This research mode can be termed *proactive* because its purpose is to develop an understanding of customer decision making sufficient to predict market behavior in a range of changed environments. The premise of this research mode is that behavior toward products and services can be analyzed in terms of factors present in an existing market

---

Reprinted from Allan B. Shocker and V. Srinivasan, "Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review," *Journal of Marketing Research* 16(May 1979):159–180, published by the American Marketing Association.

The authors acknowledge two anonymous reviewers who suggested many substantive improvements to drafts of this article.

environment. Consequently, if some future environment can be depicted in terms of levels of combinations of those factors, behavior under the changed circumstances can be predicted. In the application of this research mode, respondents primarily react to alternatives, yet provide data useful for analyzing the structure of their decisions.

More conventional research has sought pragmatically to obtain *evaluations* of alternative proposals rather than to model the basis for those evaluations. Consequently, its results are relevant mainly to those proposals specifically evaluated. Should these be modified or new alternatives arise, such research must often be repeated. In contrast, though proactive research can be used to evaluate the original proposals, it permits the researcher to predict reactions to proposals for which data were not explicitly collected, better understand customer segments and their desires, and recognize opportunities not covered by the original propositions (e.g., design new, improved strategies). Importantly, over time, research results from similar sets of customers can permit the monitoring of change in the marketplace, systematic learning by the firm, and a better, more realistic formulation of organizational market performance aspirations.

The authors contrast and critique several analytical approaches to the related problems of new product concept evaluation and generation. The purpose is not only to summarize a growing amount of research which bears on these problems, but also to distill the better methodological approaches so that the validity of the proactive idea can be improved for these and other applications. The better methodological approach in any particular application often depends on factors unique to that application or must await more decisive empirical results. Nonetheless, enough research has accumulated to warrant bringing it together and assessing its implications. The authors focus on product concept evaluation and idea generation (because this is mainly where applications have been reported) but the centrality of the "product" to marketing planning and the pervasiveness of the notion of "product concept" suggest much broader applicability for these ideas, e.g., to pricing (Srinivasan 1976) and promotion decisions (Wright 1973) and marketing decisions in the service and not-for-profit sectors (Hauser and Urban 1977; Parker and Srinivasan 1976).

Concept evaluation and generation are closely interrelated in the sense that the process of understanding customer criteria for evaluating alternatives increases the marketer's ability to generate alternatives favorably evaluated by those criteria (Shocker, Gensch, and Simon 1969). The approaches reviewed here generally fit within a framework where products or brands are represented abstractly in terms of attributes potentially relevant to customer choice. Competing brands are indicated by their perceived locations (coordinates or levels) in a "perceptual product space" whose axes are relatable to such attributes. The terms "products" and "brands" are used interchangeably because, *a priori,* it is not known whether a single

product class (by more conventional definition) or candidates drawn from different classes will constitute the set of competing alternatives (Day, Shocker, and Srivastava 1978). The approaches model customer preferences for (or choices among) these brands in terms of multiattribute models whose parameters vary for different individuals or market segments under different conditions of intended usage. The *evaluation* of a prespecified new product concept within this framework can be carried on by aggregating predicted choices of each customer or market segment over all relevant usages. This framework can potentially also be used to generate *plausible* new concepts through subjective judgment or systematic search of feasible regions of the perceptual product space for those concepts which better meet organizational performance aspirations.

The foregoing framework draws together much recent research in marketing and consumer behavior: research on the effects of situational influences on buyer behavior (Belk, 1975; Berkowitz, Ginter, and Talarzyk 1975; Hansen 1976; Lutz and Kakkar 1975; Miller and Ginter 1977; Srivastava, Shocker, and Day 1978), the definition of competitive product-market boundaries (Day, Shocker, and Srivastava 1978; Srivastava, Shocker, and Day 1978), the identification of determinant buying attributes (Alpert 1971; Myers 1970; Wilkie and Weinreich 1973), the creation of perceptual product spaces (Green and Rao 1972; Howard and Sheth 1969; Johnson 1970; Klahr 1970; Morgan and Purnell 1969), the modeling of customer decision making (Ahtola 1975; Bettman 1971; Deutscher and Ryans 1975; Hansen 1976; Hauser and Shugan 1977; Wilkie and Pessemier 1973; Wright 1975; Wright and Barbour 1977), the identification of benefit segments (Green, Wind, and Jain 1972; Haley 1968; Wind 1973), the prediction of multiattribute decisions (Berkowitz, Ginter, and Talarzyk 1975; Braun and Srinivasan 1975; Dawes and Corrigan 1974; Deutscher and Ryans 1975; Green and Wind 1973; Hauser and Koppelman 1979; Hauser and Shugan 1977; Huber 1974; Huber 1975; Jain et al. 1978; McFadden 1970; Nakanishi, Cooper, and Kassarjian 1974; Parker and Srinivasan 1976; Pekelman and Sen 1974; Pessemier and Root 1973; Punj and Staelin 1978; Ryans 1974; Wildt and Bruno 1974; Wilie and Pessemier 1973), product positioning (Albers and Brockhoff 1977; Brown et al. 1968; Haley 1968; Johnson 1971; Urban 1975; Wind 1973; Zufryden 1976), and analytical search for optimal product concepts (Albers and Brockhoff 1977; Pessemier 1975; Shocker, May, and Sudharshan 1978; Shocker and Srinivasan 1974; Zufryden 1976). A plausible theoretical basis and further empirical support are provided by the growing body of literature on the Fishbein and extended Fishbein attitude models (Ahtola 1975; Beckwith and Lehmann 1975; Fishbein and Ajzen 1975; Hansen 1976; Ryan and Bonfield 1975; Wilkie and Pessemier 1973).

The authors' concerns are restricted to analytic approaches which provide for extensive customer inputs and are potentially usable for both con-

cept evaluation and generation or product refinement (i.e., altering or modifying an existing design to improve its market acceptance). First, brief consideration is given to several approaches which do not explicitly model individual customer decision making, but rather rely on the decisions themselves and on aggregate market responses. The critique of these approaches is used to build an appreciation for the value of analyzing decisions at individual or segment levels. Next, an explicit multistage analytic framework for the proactive approach is offered and several alternative methods for implementing each stage are examined critically. Implementation results are then documented. In conclusion, applications, issues, and research needed for further development of the proactive mode of research are discussed.

Analytical Approaches for Concept Evaluation and Generation

Methods Based on Analysis of Aggregate Responses

Kuehn and Day (1962; Day 1968) and *Benson* (1965) presented approaches to infer a distribution of user "ideal points" for different levels of single product attributes. Because their procedures could be replicated for different attributes independently, they were in this limited sense multidimensional. Kuehn and Day postulated a probabilistic model to explain preferences between different attribute levels in paired comparison testing. This model enabled them to estimate the proportion of customers who had their most preferred alternative at each discrete interval of the attribute continuum. Their procedures were sufficiently flexible to permit a variety of "shapes" for this probability distribution. In contrast, Benson prespecified the functional form for the distribution and calibrated its parameters from actual preferences for a small number of contrived brands representing different levels of the attribute. He later discussed a method for fitting a distribution to actual purchase data (constructing a distribution of "choices" rather than "highest preferred levels") and extending it to consider two attributes simultaneously (Benson 1966). These researchers presented empirical results illustrating applications for their methodologies (but not testing their validity).

The contributions of these two early approaches were substantial. They provided a key research direction by relating consumer preferences to levels of product attributes. Kuehn and Day, particularly, were among the first to introduce to marketing the concepts of a most preferred ("ideal") attribute level and preferences related to distances from that ideal. Their procedures

still have utility for cases in which it is difficult to obtain many judgments from the same individual—for example, when taste is an important attribute, as with ice creams, coffees, beers, etc., satiation or boredom or a dulling of the senses may be operative (Srinivasan 1975). It is then necessary to aggregate data from many customers to estimate the preference distribution. Importantly, these researchers recognized that by comparing attribute locations of existing brands with locations of preferred attribute levels, one could discover opportunities for new brands or better understand the performance of existing ones.

Critique. Though the contributions of these approaches were substantial, so were their limitations. Most important, the models were unable *simultaneously* to incorporate the effects of multiple attributes on demand. If individuals make tradeoffs among several attributes, then an approach which optimizes the level of a product on each attribute independently may fail to lead to an optimal product design. (Though the procedures can be generalized for multidimensional estimation of consumer preference distributions (Srinivasan 1975), attribute saliences must be assumed identical over individuals as the price paid for requiring very few judgments per person). Further, the approaches were applicable only for those product categories in which the "most preferred" attribute levels would be distributed over a *finite range* of feasible levels. Some attributes such as durability or strength might, *ceteris paribus,* be thought of as "the more the better," implying infinite "ideal" levels. Other attributes might have their ideal levels at finite but infeasible points e.g., the level cannot be achieved with current technology or at reasonable cost). Still other attributes may be polytomous (e.g., American vs. Japanese vs. European automobiles, filter vs. nonfilter cigarettes) and hence a continuous preference distribution may be inappropriate. Finally, it is important to recognize that the formulations permitted individual differences only in ideal points and not in the tradeoffs between the attributes deemed salient. Though there might be pragmatic reasons (e.g., very few paired comparisons required per individual) why one may wish to ignore individual differences in tradeoffs, any procedure which forces one to ignore them under all circumstances seems undesirable.

Stefflre (1971) offered major improvements over these earlier approaches. He recognized the importance of starting with customer views of competing alternatives. After generating extensive lists of products/brands and the uses for them germane to a broad market definition, he simultaneously clustered uses in terms of appropriate products and products in terms of appropriate uses to arrive at a narrower subjective definition of the relevant product-market. He was the first to offer a customer-based analytical procedure explicitly to facilitate new product design. And he recognized the importance of customer inputs to all stages of analysis: customer protocols were analyzed for vocabulary to use in describing poten-

tial new product concepts; data on perceived similarities of products were used to construct a spatial configuration of the existing product market; preference data on existing and potential new products were explicitly incorporated into the analysis; and extensive customer inputs were used in creating physical renditions of the preferred new product descriptions.

Stefflre was among the first to apply nonmetric multidimensional scaling (MDS) of similarity judgments (Green and Carmone 1970; Green and Rao 1972) to a marketing problem. He organized his approach in several stages so that the conclusions drawn from analysis at one stage might be validated by the next. He remains one of very few researchers to be concerned with issues of testing and validation and has presented empirical evidence in support of his procedures (Stefflre 1971, 1972). Stefflre used brand descriptions to aid understanding of the determinants of product perception and preference. He believed perceptual spaces were discontinuous and used the positions of existing products and prespecified concepts to label regions. His approach facilitated the researcher's ingenuity in considering perceptual and preference data and combining them with other sources of ideas to identify new product concepts possibly very different from existing alternatives. He suggested that a range of strategic marketing decision areas could be aided by his analytical framework. For example, competitive product moves could be incorporated into the analytical structure and (with additional data collection) the structure used to explore the effects on demand of various plausible countermoves by the firm (Brown et al. 1968). Finally, to the authors' knowledge Stefflre is the only researcher to have attacked the problem of building real products and marketing plans to correspond to the desired concepts and to have created a number of new products which were actually marketed (Stefflre 1971).

Critique. There are several concerns with Stefflre's procedures. First, MDS, which he used to develop a geometrical configuration of the existing product-market, is based on perceived similarities data. The relation of such a configuration to an understanding of which concept alternatives would be *preferred* is unclear. For example, a gap may be present in such a configuration because of limited demand rather than as an indicant of untapped opportunity. Second, because Stefflre's procedures for generating product concepts relied heavily on the linguistic and analytical skill and ingenuity of the analyst, no "optimal concept" would necessarily result. Third, the improved concepts which emerged from the analyst's ingenuity must be evaluated (in terms of "drawing power") by a new data collection effort. Fourth, Stefflre aggregated individual perceptual and preference data prior to analysis and, consequently, failed to consider individual or market segment differences explicitly. Finally, Stefflre's basic premise was that when the description of a brand closely matched the description an individual would give for that brand by himself, then the individual would react to that description as he would to the brand itself. In practice, however, Stefflre

often used simple, relatively brief descriptions (of possibly complex concepts) and their validity for his intended purposes is questionable. Moreover, the premise itself has not been extensively tested.

Wind (1973) developed a multistage procedure for concept evaluation, but recognized its potential for improving product design. He sought to identify benefit segments and to determine which of several prespecified concepts was more promising (overall and for various usage occasions) for his entire sample as well as for each segment. Several evaluative measures for each concept were obtained from respondents and different analyses were used to assess the importance of various product attributes to each evaluation. Further, analysis was used to distinguish benefit segments in terms of various demographic, socioeconomic, and psychographic correlates. Finally, assessment of each concept's market position was obtained from MDS analyses of rating/rankings of concepts and other brands in terms of overall preference, appropriateness for various usage occasions, and degree of achievement of desired product benefits.

Wind recognized the utility of his procedures for product design in two ways: through conjoint measurement to estimate the part-worths of concept elements (thereby suggesting concepts or combinations of elements likely to be more favorably evaluated) and through modification of the initial concepts by judgments based on his analysis (i.e., understanding a rationale for customer decisions). Like Stefflre, he recognized the existence of submarkets (usage situations) for which a given concept might be appropriate and developed means to aggregate demand over these submarkets. Unlike Stefflre, he incorporated segment differences by identifying benefit segments for which separate analyses were later undertaken.

Critique. Unfortunately, Wind chose to present his procedures largely in descriptive terms, leaving the reader to judge which analytical technique was appropriate and which particular measures were to be collected in any specific application. As virtually every multivariate technique had a place in the framework, this task was very difficult for the user. Further, he described the creation of benefit segments *prior to* defining usage occasions. Logically, benefit segmentation might better be carried out separately for each usage submarket (rather than the reverse) because this would allow different benefits (and possible products/brands) to be emphasized by customers as a consequence of intended usage.

Myers (1976) developed benefit structure analysis (BSA) for the purpose of "finding new product opportunities in very *broad* product/service categories" (p. 23) and providing guidance for the modification or repositioning of current offerings. It had less usefulness for concept evaluation per se. A large data base was established consisting of: tasks involved in product usage within a broad category, products and brands used, benefits sought, physical characteristics of the products used, and other information related to usage (e.g., other equipment needed, involvement of other people

in the task). Measures also were sought of the extent to which each benefit and product characteristic was desired for each usage and the extent to which they were received. Myers' approach consisted of several suggested analyses based on aggregates of these data. He sought to distinguish his procedures from several of those discussed in the next section by focusing on benefit deficiencies to guide development of new products to solve problems customers have with established ones.

Critique. Myers discussed the applicability of BSA to a large number of marketing questions but provided relatively little substantiation for his claims. Most of his proposed analyses were bivariate comparisons based on aggregated measures, which obscure both the probable multivariate nature of the true relationship and the nature of underlying individual differences. Consequently, the predictive validity of such analyses can be questioned. Moreover, the sample sizes used by Myers (e.g., 500 nationally) may make his separate analyses for finer subgroups of the original sample statistically unreliable.

Myers based several of his analyses on "benefit deficiencies" defined as the differences between "benefits wanted" and "benefits received," each measured on a coarse four-point scale. Consequently, the reliability of the scaled differences is suspect. Finally, the assumption of BSA that removal of a benefit deficiency would always produce greater acceptance of the resulting product is also suspect. The interdependencies among different benefits may make it infeasible to improve one without simultaneous and possibly negative effects on others (e.g., costs of the product may increase). It is precisely this reason which makes it necessary to consider the combination of benefits and costs that a proposed new product represents to potential customers.

The approaches discussed in the next section explicitly model individual customer or segment decisions by considering the tradeoffs that are an inevitable part of choosing among feasible multiattributed alternatives. Models of individual/segment decisions are essential for forecasting the extent of market acceptance of potential new concepts for generating "optimal" new concepts. Concept generation requires the rapid, systematic evaluation of very large numbers of potential new product alternatives to determine which ones better meet the firm's objectives.

The Customer Decision Model Approaches

Though each of the following approaches is individually distinctive, they have in common a concern with modeling individual customer or market segment decisions in a multiattribute framework. It is in this sense that they are examples of proactive research. Differences between approaches are better considered within a framework which emphasizes their similarities.

1. *Determination of relevant product-markets.* Submarkets are defined in terms of the major (types of) usage situations for which concepts are to

be developed or evaluated. A set of relevant existing products/brands and customer subpopulations are identified as appropriate for each submarket. Normally there will be considerable overlap between submarkets in the sets of products deemed appropriate and the customers for whom the usage situations are relevant.

2. *Identification of determinant attributes.* For each submarket (usage situation) the set of attributes which are probable determinants of brand choice is identified.
3. *Creation of an abstract representation of each submarket.* Existing brands and descriptions of brands[1] appropriate to each submarket are represented abstractly as locations in a perceptual product space (i.e., a geometric configuration which is postulated to represent market perceptions of the position of competitive brands in terms of the determinant attributes).
4. *Development of models of individual behavior toward new concepts and existing brands.* A model is developed to predict the preferences (choice) of each individual (in a sample of individuals representative of the subpopulation of customers) from among the competing alternatives in each submarket. Though the structure of the model is assumed the same for all individuals, its parameters are permitted to vary across individuals. In some of the approaches these parameters are estimated at the market segment rather than the individual level.

5a. *Evaluation of a prespecified new product concept.* Each new concept to be evaluated is first positioned in the perceptual space corresponding to each submarket and then evaluated in terms of its differential appeal to each individual (or segment). An overall evaluation is made by aggregating choices over individuals (segments) and submarkets.

5b. *Search of the perceptual product space to generate "optimal" new concepts.* Generation of new product concepts is achieved by searching the perceptual product space for concepts which "optimally" achieve desired marketing objectives. This step is accomplished by evaluating market potentials for the new concepts (forecasted by the procedures used in stage 5a) together with their cost and competitive implications and considering the effects of cannibalization on the current product line of the firm.

The approaches of Johnson (1971), Lehmann (1971), Shocker and Srinivasan (1974), Rao and Soutar (1975), Hustad, Mayer, and Whipple (1975), Pessemier (1975),[2] Urban (1975), and Hauser and Urban (1977) are reviewed here. They differ in terms of their relative concern with the different stages, the specific means by which each stage is operationalized, and their focus on concept evaluation versus generation. Table 2-1 summarizes each approach and is the authors' only effort to provide an overview of each. In the discussion following, concern is with contrasting alternative ways of implementing each of the stages (i.e., which method is better for each stage and under what circumstances) and not with the different

researchers' approaches per se. In other words, the emphasis is on the rows of table 2–1 rather than the columns.

The foregoing framework is primarily useful for locating "new" product opportunities which may not be substantially different from current alternatives. (The vast majority of "new" products which enter the marketplace tend to be of this kind.) This is a natural consequence of starting from the structure of existing markets (although a broad market definition in the manner of Stefflre and Myers will always be advantageous). By creative use of descriptions of fictitious product concepts it may be possible to extend the framework to analyze a broader set of "new" product possibilities. Such fictitious alternatives might have some attributes not available in existing ones or have levels of attributes lying outside the ranges of existing ones.

1. Determination of relevant product-markets. At this stage the purpose of analysis is to identify a set of competitive product/brand alternatives and their potential customers. Many of the approaches (Johnson 1971; Lehmann 1971; Rao and Soutar 1975) are based on the assumption that products and customers composing a relevant product-market are given or otherwise apparent. Both Urban (1975) and Hauser and Urban (1977), however, determine the set of relevant brands from customers' evoked sets—operationalized as brands last used, ever used, on hand, or would not be considered. These sets are elicited from subjects presumably under some prespecified broad market definition, for example, household cleaning agents, duplication processes (copiers). The product-market would be defined as the totality of brands elicited from some prespecified minimum proportion of respondents. This procedure provides a "familiarity" definition of that market. In contrast, Pessemier (1975) suggested managerial judgment to determine relevant assortments, but would temper that judgment by considering products which are "natural substitutes" and which the user is able to purchase without extraordinary effort or budgetary strain. He would exclude products which represent minor competition. The logic of the latter criterion is unclear in view of the purposes of later analysis (to identify determinant attributes) and the dynamic nature of markets (small share brands may over time become large and vice versa).

The primary weakness of the foregoing product market definitions is that they fail to consider that an overall market may consist of submarkets which behave very differently. Recent research in marketing and social psychology has shown that situational effects are important moderators of customer behavior (Belk 1975; Berkowitz, Ginter, and Talarzyk 1975; Lutz and Kakkar 1975; Miller and Ginter 1977; Srivastava, Shocker, and Day 1978). Day, Shocker, and Srivastava (1978) recently reviewed and critiqued several alternative customer-based approaches and argued for the determination of situationally defined submarkets (in terms of intended use) within a broad product-market definition. Any given product could be a

part of more than one such submarket, but the submarkets themselves would each contain distinct alternatives.

The approaches which incorporate situational effects (Hustad, Mayer, and Whipple 1975; Shocker and Srinivasan 1974) have benefited greatly from the earlier work of Stefflre. Recall that his approach began broadly by generating long lists of products and uses through unstructured interviews (e.g., focus groups). Respondents then were asked to indicate which of the products they considered appropriate for each use. Aggregation of these data and then simultaneous clusterings of uses (in terms of similar products considered appropriate) and products (in terms of appropriateness to similar uses) were used to define submarkets. Customer (or, less preferably, managerial) judgment could indicate which of the situation-product clusters (submarkets) to retain (e.g., a sample of customers could be asked to indicate the relative frequency with which different situations occur). Those individuals would be the potential customers for that submarket. One desirable aspect of Stefflre's approach was his consideration of "home remedies" as well as commercial products in developing a list of product alternatives. The insights that might be gained from such noncommercial preparations and outlier products may increase the likelihood that the framework will lead to relatively novel new concepts.

Belk (1975) and Srivastava, Shocker, and Day (1978) attempted to create a situational typology from factors which appear to "explain" clusters of objective situations. This typology resembles a factorial experimental design. A number of objective situations could be associated with each given cell (situational type) and products judged appropriate to each cluster of situations (by some prespecified minimal proportion of potential customers) would define the product market. The usefulness of situational definitions of submarkets is very much conditional upon being able to define a relatively small number of situational types (otherwise the amount of data needed would be overwhelming). Using the typology, one needs to investigate only the cells containing a number of frequently occurring or "important" situations.

2. Identification of determinant attributes. Determinant attributes are those which both distinguish the product alternatives in the relevant market and can be reliably associated with customer choice (or preference)(Alpert 1971; Myers 1970). Such attributes are, in principle, relatable to the benefits/costs customers seek through their purchase and use. They may be psychological (or sociological) as well as physical. Because they may be idiosyncratic, the attributes relevant for further analysis are those used by some prespecified minimal proportion of potential customers. (Zero importance weights in the individual decision models discussed at stage four would indicate that particular attributes were not salient for the individual.)

A customer-based methodology requires such attributes be specified in terms meaningful to customers. The marketer's concerns, however, are with

Table 2-1
Summary of Methods Used to Operationalize Stages in Concept Evaluation/Generation Framework

Stages/ Approaches	*Johnson (1971)*	*Lehmann (1971)*	*Shocker Srinivasan (1974)*	*Rao and Soutar (1975)*
1. Determination of relevant product markets	Competitive products assumed known	Not discussed	(Stefflre's) Products-by-uses analysis	Not discussed
2. Identification of determinant attributes	Direct questioning	Direct questioning Executive judgment MDS—multidimensional scaling of similarity judgments	Kelly's repertory grid MDS Direct questioning Regression	Functional attributes determined judgmentally
3. Creation of perceptual product space	Multiple-discriminant analysis	Direct measurement of attribute levels	Direct measurement of attribute levels	Direct measurement of attribute levels
4. Modeling individual or segment decision making	Interval-scaled preferences Direct scaling of ideal points Regression to estimate attribute weights Choice not modeled	Direct scaling Probabilistic choice model	Ordinal preferences LINMAP First-choice model Probabilistic choice model (Luce)	Interval-scaled preferences using Thurstone's law Regression Probabilistic choice model (Luce)
5. Evaluation of/search for new-product concepts	Heuristic—locate new products near concentrations of ideal points	Subjective—generate arbitrary new concepts and use framework to predict market share/present value of sales	Nonlinear program Heuristic search through coarse and fine grids Modified gradient search Costs explicitly considered	Nonlinear program
6. Remarks		Based on expectancy-value theory	First overall approach to recognize importance of all steps in conceptual framework	Uses only functional product attributes

Table 2–1 continued

Hustad et al. (1975)	*Pessemier (1975)*	*Urban (1975)*	*Hauser and Urban (1977)*	*Other Approaches*
Ad hoc situations used Evoked set of products	Evoked set of products	Evoked set of products	Evoked set of products	Three-mode factor analysis (Belk 1975) and principal-components analysis (Srivastava et al. 1978) to define situational typologies Review of methods (Day et al. 1978)
Direct questioning Judgment Hustad's procedure to determine optimal number of attributes	No specific techniques suggested	Direct questioning Kelly's repertory grid MDS	Direct questioning Kelly's repertory grid MDS	Review of methods (Alpert 1971; Myers 1970; Wilkie and Weinreich 1973)
MDS	Multiple-discriminant analysis	Factor analysis MDS	Factor analysis MDS	MDPREF (Green and Carmone 1970; Green and Rao 1972)
Ordinal preferences PREFMAP Choice not modeled	Ratio (dollar metric) preferences PREFMAP Probalistic choice model (Luce)	Ordinal preferences PREFMAP or LINMAP Prediction of long-run market share	Ordinal preferences PREFMAP or LINMAP Utility theory Probabilistic choice model (multinominal logit)	Review of methods (Green and Srinivasan 1978; Green and Wind 1973) Multiplicative models (Nakanishi et al. 1974) Conjunctive, disjunctive models (Wright 1975) Choice models (McFadden 1970; Pessemier et al. 1971; Punj and Staelin 1978) Brand-specific effects (Srinivasan 1976)
Not discussed	Heuristic search (STRATOP) gradient procedure Costs explicitly considered	Subjective (product refinement)	Subjective (product refinement)	Nonlinear programming (Albers and Brockhoff 1977; Zufryden 1976) Gradient search (Morgan and Purnell 1969)
First approach explicitly using situations	Mentions need to consider competitive reactions to moves by the firm	Uses multistage data collection to predict ultimate trial and brand-switching behaviors and long-term market share	Models manager's decision process as well as customer's and integrates the two	

things he/she knows how to influence or control—physical properties, price, promotion frequency and message, etc. Shocker and Srinivasan (1974) termed attributes meaningful to both customers and marketers as actionable. Objective attributes such as price should be directly actionable. Many subjective attributes can be made more actionable by developing anchored rating scales. Actionability need not require that customers' and marketer's sets of attributes be identical, merely that there be psychophysical transforms between the two (e.g., levels of sweetness corresponding to levels of sugar contents, levels of wear resistance corresponding to fiber content or weave of material). Green (1975) expressed doubt that such transforms can, in general, be established. Fortunately, the problem is mitigated if the purpose of analysis is concept evaluation. Then it is only necessary that present product alternatives and new concepts be positionable within each relevant submarket. But when concept generation by some form of search (step 5b) is the objective, actionability is an important concern because the manufacturer must develop an "extended product" corresponding to the desired concept. The analyst must either model customer decision making as a function of actionable attributes alone, develop the appropriate psychophysical transforms, or find surrogate means (e.g., by locating actionable descriptions in the perceptual space in an attempt, like Stefflre's to label regions of that space). These difficulties may be less severe in industrial marketing contexts where buyers are relatively sophisticated, but otherwise they could limit applications.

Several alternate means are available for identifying determinant attributes. Direct questioning of customers, Kelly's repertory grid, a measure consisting of ratings of attribute importance multiplied by the dispersion of products on that attribute (Wilkie and Weinreich 1973), multivariate analyses of the relation between product attribute ratings and measures of preference or choice, and projective techniques such as word associations are among those which have been suggested (Shocker and Srinivasan 1974). These still need to be investigated for their relative efficacy in the present context. A less direct approach is MDS of perceived similarities between brands to create a geometric representation of brands to assist the analyst in recognizing criteria that underlie perceived similarity. Such indirect means may prove advantageous as respondents would be asked for behaviors that are more natural for them (e.g., judged similarity), and thus the data they provide might be more valid. It is important to realize that attributes identified by such indirect methods may merely distinguish the product alternatives rather than explain differences in preference among them. Hence additional analyses (e.g., multiple regression) may be needed to establish a relation to preference. Perhaps the nonmetric scaling of preference rather than similarities data would provide this more direct relation (Green and Carmone 1970; Green and Rao 1972). A concern with nonmetric pro-

cedures, however, is that the evoked sets of existing brands in many product-markets are too small or too varied across individuals for the nonmetric techniques to achieve reliable results (Hauser and Urban 1977).

3. Creation of a perceptual product space. The various approaches differ in the methods suggested for creating perceptual product spaces. They are similar, however (except for Urban's), in their assumption that different market segments (or individuals) have a common perceptual framework. Stefflre (1972) presented empirical evidence suggesting considerable agreement among judges in their perceptions (correlations in the range of 0.75 to 0.92). In examining perceived similarity of cigarette brands, Klahr (1970) also reported highly significant correlations between judges. Day, Deutscher, and Ryans (1976) also found the assumption of homogeneous perceptions reasonable in several markets studied. However, the hypothesis of homogeneity of perceived similarities was rejected by Ritchie (1974) in his study of leisure activities. For the purpose of evaluating existing concepts, homogeneity of perceptions is not essential because it may be possible to map the concept to be evaluated into each individual perceptual framework. However, for the purpose of concept generation, the assumption of a common framework is obligatory. This is because evaluation and/or search for desirable new concepts requires that it be feasible to evaluate large numbers of candidates efficiently and to translate any chosen one into specific implications for marketing plans, e.g., physical product characteristics, promotional message, packaging (Brown et al. 1968; Stefflre 1971). Such tasks are difficult, and would be virtually impossible were it necessary to coordinate multiple customer spaces with a single manufacturer's space. Finally, it is important to realize that some individual differences in perception are captured effectively by the attribute salience weights in the models of customer decision making developed in stage 4. These weights effectively permit individual "points of view" through idiosyncratic shifting of the origin and differential stretching of the axes of the perceptual space. Thus, perceptual homogeneity need only be presumed up to a linear transformation (i.e., the relative positions of the brands on each attribute are presumed the same for different individuals). Procedures such as INDSCAL, which involve this assumption, have been shown to have substantial validity (Green and Rao 1972; Kinnear and Taylor 1973). In fact, Ritchie (1974), who rejected the hypothesis of homogeneity of perceived similarities, found meaningful results using INDSCAL.

Except when MDS is used, creation of a perceptual product space typiclly begins with customers scaling existing product/brands (and perhaps a set of fictitious brands) independently on each determinant attribute. Bipolar adjectival scales, e.g., seven-point scales such as high-low, modern-old fashioned (Johnson 1971; Lehmann 1971; Pessemier 1975; Shocker and Srinivasan 1974), agree-disagree scales, i.e., where perception is measured

by strength of agreement with statements such as "I would be paying too much for this service" (Hauser and Urban 1977), or anchored scales (Wind 1973) are typically used for this purpose. The prespecification of bipolar or agree-disagree scales implies the attribute is a continuum and that the psychological distance between extreme positions is divisible into a fixed number of approximately equal intervals. Some attributes may not be readily representable by such scales (e.g., nominally scaled attributes such as "country of origin" for a car or "existence of a product feature" as in filter vs. nonfilter cigarettes). Others can (often with redefinition) be representable by either type of scale (e.g., durability, warranty coverage). It is desirable that analytical procedures in the subsequent stages have the flexibility to cope with such scaling diversity.

Some researchers (Lehmann 1971; Rao and Soutar 1975; Shocker and Srinivasan 1974) suggest working directly with respondents' attribute perceptions (direct scaling). Each product then can be located at its modal (mean) perceptual value (across judges) on each attribute scale. Other researchers suggest multiple discriminant analysis (Johnson 1971; Pessemier 1975; Pessemier and Root 1973) or factor analysis (Hauser and Urban 1977; Howard and Sheth 1969; Morgan and Purnell 1969; Urban 1975) of these attribute data or MDS (Urban 1975) to determine a configuration in reduced space. As discussed by Johnson, the multiple discriminant procedure involves analysis of a three-way data matrix of products by attributes by respondents to find a small number of weighted linear combinations of those attributes which discriminate "best" among the products (represented by different individuals' perceptions). Average attribute scores are used to position products in the space spanned by these discriminant functions (Johnson 1970). Factor analysis procedures analyze a two-way matrix of brand ratings by attributes (averaged over individuals) to find a reduced space configuration spanned by dimensions (factors) subjectively interpretable by using the factor loadings of the original attributes. Products/brands are represented in this perceptual map by their factor scores.

Discussion. Each methodological approach has certain advantages and disadvantages. Only direct scaling permits the analyst to "know" what dimensions underlie the perceptual product space. Further, locations of particular products in the structure are not affected by the number or locations of other products (i.e., the configuration is stable in the sense that products or brands normally can be added to an existing structure without altering the relationships between brands previously positioned). Neither of these conditions generally applies when a configuration is determined by reduced space approaches. Whether such differences are consequential is the important empirical question. Direct scaling has a major disadvantage in that it typically leads to larger numbers of dimensions. This feature in turn may preclude reliable estimation of parameters in the customer behavior models on the basis of preferences for existing products alone. Instead,

parameter estimation may have to be done by use of a set of fictitious brand descriptions (Hauser and Urban 1977; Parker and Srinivasan 1976; Wind 1973). Descriptions may be desirable in any event because actual brands are often not sufficiently distinctive to provide reliable estimates of tradeoffs (see note 1). As already noted, for concept evaluation it is necessary only to be able to map a concept into a unique location in the space. Multiple discriminant analysis and factor analysis appear to be adequate methodologies for this purpose because the dimensions are directly relatable to the prespecified attributes (whereas MDS requires judgment on a criterion of overall similarity). Johnson (1970, 1971) noted that tests of statistical significance are available for each added dimension, that these methodologies are well known in the literature, and that solutions (configurations) are unique (conditions which do not hold for MDS). Compared with other reduced space approaches, MDS requires no prespecification of attributes. Consequently, its use could serve as a partial check on the set of determinant attributes chosen at stage 2. Subsequent analysis could be based on factor or discriminant analysis. In the reduced space approaches, fewer parameters need to be estimated in operationalizing individual (segment) preference models (discussed in the next section). This feature may make individual level analysis more feasible and may increase reliability of measurement.

There are, however, several limitations to reduced space approaches. Dimensions underlying the reduced space may not be interpretable in actionable (or any other) terms. Consequently, these approaches are less applicable for concept generation because a point in the reduced space corresponds to an infinity of points in the original space of determinant attributes. Further, unless care is taken in defining the underlying attributes in which products are originally scaled (prior to factor or discriminant analysis), problems arise in the modeling of preference judgments (at stage 4). For example, suppose the product category is automobiles. Two attributes which may be salient to preferences are "size of car" and "operating costs." These attributes, for technological reasons, normally will be strongly and positively associated and hence are likely to load in the same direction on the same factor (or contribute similarly to each discriminant function). Preferences for some individuals could reasonably be expected to be related positively to "size of car" but *negatively* to "operating costs." Consequently, the importance of the reduced space dimension for preference is unclear because of these contrary effects. The crux of this argument is the question of whether preference is really a function of the original attributes or of "latent" attributes potentially discoverable through a reduced space method. If it is a function of the original attributes, care must be taken in scaling them so that they will load on the reduced space dimension in the same way they affect preferences—a condition not always feasible in practice.

Hauser and Koppelman (1979) conducted an empirical investigation of several alternative methods for determining perceptual spaces and developing models of preference, separately and in combination. They compared all the major methods on such criteria as ability to provide managerial interpretation, to predict consumer preferences accurately, ease of use, and cost of implementation. Factor analysis was found to be superior to MDS on all criteria and superior to discriminant analysis on interpretability and predictability but equal on the other two criteria. The directly scaled attributes were judged best on ease of use, but poorest on managerial interpretability [they presented "too much information to readily internalize for strategy development" (1979, p. 16)]. Because of high multicollinearity in the raw attribute data, direct scaling was judged less suitable for use with *statistical* preference models (e.g., regression) but otherwise gave good predictability. Cost of direct scaling was generally the least, except in conjunction with the logit model of preference (see stage 4) where costs increased rapidly with number of variables. Though these results are based on a single empirical data set, they seem to suggest advantages for factor analysis and direct scaling.

4. Development of models of customer behavior. The behavior that most approaches model is individual (or segment) *preference* decision making . Preferences may be linked subsequently to market behavior (choice) by an independent modeling step. The primary reasons for modeling preferences separately are that more information about *decision making* is obtainable (an entire preference ordering is obtainable whereas choice provides only the decision itself) and that such data are easily and reliably collected. Moreover, choice may be affected to a greater extent by factors not readily controllable by the marketer and hence must be modeled probabilistically.

Customer decision making is typically modeled by use of one of several versions of what is termed a compensatory model (Green and Srinivasan 1978; Green and Wind 1973). (Hauser and Urban's utility theory approach is the exception.) *The "ideal point" version* represents the decision maker by a location in the perceptual product space termed his "ideal point." The individual's utility for a given brand is related inversely to its weighted Euclidean distance from his ideal point. The individual is presumed to prefer alternatives "closer" to his ideal point over those "farther away." The "ideal point" is merely a reference location in the perceptual space from which distance measures can be taken and need not correspond to any real or feasible product/concept. *The "vector version"* models utility directly as a weighted sum of the attribute levels of each alternative. In both models, if dimensional units of measurement are standardized, the weights may be interpretable as measures of the relative importance (salience) the customer attaches to each attribute dimension in making tradeoffs. The vector is a special case of the ideal point version in which levels for every attri-

bute dimension are plus or minus infinity so that the individual is presumed always to prefer more (less) of each attribute. The *conjoint measurement version* considers a finite number of levels for each attribute and permits a preference function to bear an arbitrary relation to attribute levels through the use of dummy variables.

Much research has examined whether consumers actually use linear compensatory rather than seemingly simpler evaluation models such as lexicographic and conjunctive (cutoff) models (Hansen 1976; Wright 1975). Generally this research has shown that some consumers appear to use each of the models, but generally prefer those requiring simpler cognitive processes. This research, however, has investigated decision making at a single point in time rather than as a process with learning over time. The compensatory model is cognitively less demanding over time where, because of its additive nature, only the effects of *changes* in the perceptual structure of the individual need be considered to modify a preference judgment arrived at previously. Comparisons can be made between two alternatives by explicitly considering only attribute *differences* which distinguish them. Further, the present framework views markets from a situational perspective where the set of competing brands and concepts are only those *considered appropriate* for the intended usage, so that tradeoffs may be more necessary. The number of alternatives to be evaluated thus should be relatively small as should the number of determinant dimensions so that a compensatory model may prove sufficiently simple. Finally, it can be noted that even if the foregoing conditions were not to hold, such a model need be only a *representation* of the decision process (i.e., have high predictive validity), not necessarily the process itself (Dawes and Corrigan 1974). As studies of protocols of customer decision making have shown, the decision processes themselves may be too complex (Bettman 1971) to model in a tractable way. The predictive ability of the compensatory model has been demonstrated empirically in several studies (Braun and Srinivasan 1975; Dawes and Corrigan 1974; Green and Wind 1973; Hauser and Shugan 1977; Huber 1975; Parker and Srinivasan 1976; Pekelman and Sen, 1974) and gives good results even when decisions are actually made using more complicated rules (Dawes and Corrigan 1974; Green and Devita 1975). Thus, the authors conclude that the compensatory model should prove adequate for most applications of the framework.

Scaling Model Constructs. Model constructs (ideal points, attribute saliences) can be measured by either direct or derived scaling. Direct scaling involves asking respondents to describe their ideal product on a prespecified attribute scale (typically the same bipolar adjectival scales used with existing products) and to scale the relative importance of each attribute on some prespecified "importance" scale. Derived scaling treat the constructs as

parameters in a preference model to be estimated by either metric or nonmetric multivariate statistical procedures. Though most authors advocate one approach or the other, Johnson (1971) favored direct scaling of ideal points but used multiple regression to estimate the importances of attributes. Hustad, Mayer, and Whipple (1975) use a method developed by Hustad to find the relative contribution and "optimal" number of attributes to predict rank order product preferences.

Several approaches have been proposed for obtaining derived estimates of model constructs. Shocker and Srinivasan proposed their LINMAP technique (Srinivasan and Shocker, 1973*a, b*) to estimate the two sets of parameters simultanously for each individual (or segment) on the basis of ordinal preferences between pairs of products. LINMAP provides for a "mixed mode" option in which, for some attributes, ideal points need not be finite if fit to the data can be improved. Thus the data determine whether the model incorporates either or both the ideal point and vector modes, if this option is desired by the analyst. (The conjoint measurement model also can be estimated by LINMAP.) Pessemier, in contrast, scaled preference judgments using his dollar metric procedures (1975; Pessemier et al. 1971). He claimed these procedures led to ratio-scaled preferences (where scale differences express differences in product value in monetary terms). These dollar metric values then served as inputs to the PREFMAP scaling algorithm which estimated ideal points and attribute saliences. Urban (1975) developed metric measures of preference (by asking respondents to split 100 points between each pair of brands in proportion to their preference) and used PREFMAP to estimate parameters using both ideal point and vector versions. He pragmatically selected whichever version better fit his data. Hauser and Urban (1977) mentioned several alternative methods for modeling preference: direct consumer statements of model parameters, statistical estimation of choice parameters (e.g., using PREFMAP), and von Neumann-Morgenstern utility theory.

Although none of the approaches specifically made use of conjoint measurement models for the estimation of parameters, these have received widespread managerial acceptance and might be profitably used. Two general methods for obtaining ordinal preference judgments have been proposed and they have implications for the predictive validity of the resulting models. In the full profile method of data collection, respondents compare product "alternatives" represented abstractly in terms of some level of *each* underlying attribute (Green and Srinivasan 1978). In tradeoff analysis respondents are presented possible pairs of attributes and are asked to rank in preference all the hypothetical concepts that could be described by combinations of levels of these two attributes (Green and Srinivasan 1978; Johnson 1974).

Discussion of Alternative Models and Methods of Scaling. The choice among ideal point, vector, and conjoint measurement versions of the com-

pensatory model should logically depend on the nature of the determinant attributes. Both infinite and finite ideal points are defensible for particular attribute definitions, e.g., one may always prefer lower cost (vector) but may prefer moderate levels of sweetness or size of automobile to extremes (ideal point). Consequently, the mixed model of Shocker and Srinivasan may have empirical advantages for the general case. Otherwise one should seek to define the attributes themselves so that the models used are appropriate *a priori* (Myers and Shocker 1978).

A major advantage for direct scaling of parameters is that it affords flexibility in accommodating different distance metrics in the ideal point version. Indirect methods of model calibration (such as PREFMAP, LINMAP, and multiple regression) must generally presume a Euclidean metric. In his empirical work, Lehmann (1971) found slight improvement in predictive validity by using the city-block metric. However, Huber's (1974) review of field studies of multiattribute decision making showed the evidence to be less conclusive. Green (1975) indicates that Euclidean approximations to other distance metrics "have been so close as to obviate the need to search for the 'correct' metric in all but the most esoteric of applications." Thus the flexibility of direct scaling with respect to difference metrics does not appear to be empirically consequential. Additionally, direct scaling of ideal points has the major disadvantage that it is not truly multidimensional (scaling of all parameter values is not done simultaneously), and thus leads to a likelihood of halo effects (Beckwith and Lehmann 1975). Further, direct scaling involves the implicit assumption that customers consciously make decisions in the manner of the model and hence can validly understand and scale model constructs. Limited empirical evidence indicates that consumers have difficulty in understanding the ideal product concept sufficiently well to scale it reliably (Myers and Chay 1977; Shocker and Srinivasan 1974).[3] Finally, if the preceding analysis incorporates a single perceptual product space for all respondents, individual differences in perception may be incorporated into the framework only through the use of derived importance weights.

Though these arguments tend to favor derived estimates, direct estimation continues to have strong adherents. One possible reason for this is the confounding of these multiattribute models with models of the Fishbein/Rosenberg attitude theories (Wilkie and Pessemier 1973), where direct estimates are always used. There are numerous similarities, e.g., the models have similar mathematical structure (Wilkie and Pessemier 1973), both have situational effects incorporated into their extended forms (Fishbein and Ajzen 1975; Ryan and Bonfield 1975), and both have been used to estimate measures of behavioral intention rather than overt purchase behavior (Fishbein and Ajzen 1975). The distinctions lie principally in the different purposes for which each was developed and in how model constructs are defined. Fishbein saw his work as a theory of attitude structure and change and hence he was primarily interested in explanation. He considered predic-

tion simply as a means for validation. Multiattribute models, in contrast, were developed largely for predictive purposes; explanation is regarded as desirable rather than essential. Constructs also are defined somewhat differently in the two approaches. Fishbein operationalized direct measures of his theoretical constructs. This was necessary to prevent tests of model structure from being confounded with measurement of the constructs. In contrast, in the multiattribute models the structure was assumed correct as specified and certain parameters (construct measures) were estimated so as to maximize goodness of fit to a criterion (e.g., preference ratings or rankings). Consequently, models using derived measures tend to have high predictive validity (Braun and Srinivasan 1975; Hansen 1976; Hauser and Koppelman 1979; Huber 1975; Jain et al. 1978; Parker and Srinivasan 1976). As ongoing research continues to validate the expectancy value model structure, derived estimates may also prove superior to direct measures of its constructs. It is thus useful to regard the expectancy value framework as a theoretical backdrop for the present approaches (Ryan and Bonfield 1975).

Consider each of the methods for obtaining derived estimates. Multiple regression desirably should fit preference data which are at least intervally scaled. But such data may be difficult to obtain reliably from simple ratings. So constant sum methods (Urban 1975), dollar metric methods (Pessemier 1975; Pessemier et al. 1971), Bradley-Terry-Luce procedures (Johnson 1970, 1971), and the Thurstone's Case V procedures (Rao and Soutar 1975; Wind 1973) have been proposed. Some of these procedures impose restrictions on the preference-generating process which may prove untenable and have not been adequately tested. The use of Thurstone's Law of Comparative Judgment is particularly questionable here. The logic of this law presumes that the population of respondents is homogeneous—a condition which seldom obtains with preference judgments unless considerable care is exercised in segmenting the population for homogeneity prior to data collection. Apart from the problem of scaling the dependent variable, conventional regression analysis may suffer because it cannot preclude negative salience parameter values for ideal point models—a fact which can complicate interpretation. However, regression is widely known and computer programs are readily available; it is very robust, which implies that even an ordinal-scaled dependent variable may produce reliable results (Cattin and Wittink 1976). If interval-scaled preferences can be assumed, the statistical properties of parameter estimates can be specified.

Because conjoint measurement models use polytomous representations of all attributes, they create possible ambiguities when attributes are really continuous, e.g., requiring interpolation between adjacent part-worths (i.e., utility values for different attribute levels) to establish the value for intermediate attribute levels. They pose further difficulties in extrapolating

preference measures for levels outside ranges in which the attributes were originally specified. Many parameter estimates (many part-worths) may be required, necessitating relatively large numbers of judgments. (If the attribute is really continuous and the part-worth utility function is monotonic and approximately linear, a single importance weight might otherwise have sufficed.) Tradeoff analysis, though the instruments are considerably easier to administer to respondents than are those based on full profiles of stimuli, is limited by the assumption that, when comparing concepts described only in terms of two attributes, the respondent holds all other attributes constant. In cases involving considerable correlation across attributes (e.g., size, gas economy, luxury, and price of automobiles) such an assumption seems untenable. On the other hand, conjoint measurement models have been widely accepted by management, due in measure to the ready interpretability of attribute part-worth functions. The tradeoff and full profile methods are compared in detail by Alpert, Betak, and Golden (1978), Green and Srinivasan (1978), Jain et al. (1978), and Montgomery, Wittink, and Glaze (1978).

Although PREFMAP is more widely known, LINMAP dominates it in several important respects without having substantial disadvantages. Both techniques use ordinal preference judgments as input. Recent research has indicated that because of local optima, the optimization procedures used in methods such as PREFMAP produce solutions only marginally different from those produced by multiple regression, thereby making LINMAP the only globally optimal nonmetric procedure for analyzing preferences (Cattin and Wittink 1976). In LINMAP, attribute importance weights can be constrained non-negative, and thus the problems of interpretability that could arise with regression or PREFMAP (e.g., negative attribute saliences) are avoided (Pessemier 1975, p. 52). Because estimation of conjoint measurement models can also be undertaken by using LINMAP (either with full profile data or that obtained through tradeoff analysis), all three versions (ideal point, vector, conjoint) can be incorporated in the same model. However, as with all nonmetric procedures such as LINMAP or PREFMAP, no statistical properties of the estimates or of the goodness-of-fit statistic for the overall model are available.

The utility theory approach of Hauser and Urban (1976) appears to show considerable promise in the few empirical applications reported. This approach follows from a rigid set of axioms and emphasis on verifiable assumptions to derive unique functional forms for the preference functions. Consequently very few data collected from each respondent (i.e., virtually no measurement error is assumed, in contrast to the statistical approaches). However, the data collection requirements are formidable. Measures of risk must be obtained by presenting respondents with lotteries (which involve choice between certain events and uncertain events while probabilities are

varied to locate indifference points). To parameterize the multiattribute utility function, explicit measures of tradeoffs between attributes must be obtained (e.g., if alternative *A* has attributes of, say, price and quality at $10 and 20 units respectively, what price would the individual pay to be indifferent between *A* and alternative *B* whose quality is 30 units?). Hauser and Urban conducted their empirical work with MIT undergraduates who had been extensively briefed prior to data collection. It is doubted that such techniques could be applied successfully to a less sophisticated population.

In one computer simulation study, Cattin and Wittink (1976) found that regression, MONANOVA (the method normally used in estimating conjoint models), LINMAP, and logit (see section on choice models) did not differ very much in their cross-validities. Regression had a slight advantage when the attribute weights did not differ markedly from each other, but LINMAP seemed to have a slight advantage when weights varied markedly across attributes. Using empirical rather than simulated data, Jain et al. (1978) found that MONANOVA, LINMAP, logit, and regression yielded roughly the same level of crossvalidity, with logit and LINMAP being the slightly preferred procedures.

In summary, then, the choice between ideal point and vector models logically should depend on specification of the determinant attributes. A mixed model appears to offer considerable flexibility. Because the major concern of the proactive approach is with prediction, derived estimates for model parameters would appear to be superior to direct measures. Overall, the different estimation methods do not seem to differ very much in their predictive validities. Metric methods (e.g., regression) and nonmetric methods (LINMAP, MONANOVA, logit) might be used together as a rough check on the robustness of the results.

Choice Models. The various approaches differ in the concern given to linking measures of preference with purchase behavior (or choice). Many simply stop with the modeling of preference behavior and treat share of first preference as an adequate criterion. Choice differs from preference in that many factors over which the marketer may have limited control may affect the purchase (e.g., number of facings, recognition of the product on the shelf, special deals, and personal recommendations). Consequently, the modeling choice has generally emphasized probabilistic models.

Lehmann (1971) argued that probability of choice should be based on relative proximity of a brand to the ideal. By clustering customers and computing average values of the model parameters for each cluster, he proposed to estimate sales for a given brand by summing (over clusters) the share of market each brand received (based on purchase quantity for the product class times the probability the brand is preferred). In his formulation, all brands have a nonzero probability of purchase by each cluster. For individ-

ual level data, Shocker and Srinivasan (1974) suggested determination of a parameter k to operationalize a consideration set for each individual consisting of the k brands closer to his ideal point. Distances to products within this set can be related inversely to their probability of choice and products outside this set can be arbitrarily assigned a zero probability of purchase. This kind of process would also provide the rationale for a zero order model of stochastic brand choice behavior (Montgomery and Urban 1969). Pessemier et al. (1971) described a similar model of choice which used dollar metric preference data (instead of distances). The dollar metric preference measure for a given brand is raised to a power and this factor is expressed in relation to the sum of similar factors for all alternatives in the individual's choice set. Pessemier et al. (1971) discuss the relation of this probability model to the Luce model of axiomatic choice. Recently, Nakanishi, Cooper, and Kassarjian (1974) proposed a similar functional form in which a measure of attitude toward purchase serves the function of Pessemier's dollar metric preference. Pessemier et al. (1971) empirically tested several models similar to those mentioned. They found excellent prediction from such models. Importantly, these researchers also compared a single choice model (where the most preferred product is always chosen) with more complex models and found this simpler model gave aggregate predictive results approximately as good. Additional empirical support for the single choice model is provided in by Braun and Srinivasan (1975), Nakanishi, Cooper, and Kassarjian (1974), and Parker and Srinivasan (1976).

Whether a single choice model or a probabilistic model of the preference-choice relationship is superior is both an empirical and a pragmatic question. The error component in choice models may be larger for "less important" purchase decisions (e.g., soft drinks rather than automobiles). Thus, less important decisions may need to be modeled probabilistically whereas more important ones might be modeled by using the single choice criterion for the same predictive validity.

Hauser and Urban make use of the multinominal logit model of McFadden (1970). This model is based on the postulate that there is a true utility which completely determines a customer's choice (i.e., a customer chooses that alternative with the largest utility), but one can only observe part of this utility. The true utility is equal to an observable part (represented by a measure of preference such as ideal point-product distance) plus a (Weibull distributed) random error. Deductive reasoning leads to the logit model. This model is calibrated by using actual choice data, e.g., based on measures of intent to purchase, choice among proxies such as descriptions (Punj and Staelin 1978), or possibly actual choice from test market results. The appeal of this model is that it explicitly models a stochastic behavior (probability a given brand is chosen) at the segment level in a manner which requires no metric assumptions about preference rankings. Its drawback is

that it usually estimates average importance weights across individuals to gain degrees of freedom. This requires that a relevant basis for creating market segments exist *a priori*. (The logit model can also be estimated at the individual level.) One of its limiting assumptions, however, is the so-called "independence from irrelevant alternatives" assumption which may restrict applicability because in many cases it would not be reasonable.[4]

Hauser and Koppelman (1979), from their empirical comparison of different techniques for modeling customers' perceptions and preferences, concluded that logit analysis and preference regression (used by PREFMAP) yield similar managerial interpretations and predict better than two other methods (i.e., unit weighted multiattribute models and expectancy-value models) when they are used with reduced space methods for generating perceptual product spaces. When the determinant attributes are used directly in creating the perceptual space, the expectancy-value and unit weight models perform well. On their other criteria (ease of use and cost) the logit model was inferior to the others. Hauser and Koppelman's findings were based on an empirical study of choice among seven shopping centers. The small number of choice alternatives precluded model calibration at the individual level—a very limiting aspect of their study. It is probable, therefore, that the performance of the logit model (the only stochastic model tested) was favored by this research setting.

Srinivasan (1976) proposed a modification of the multiattribute model which could be used to predict first choice. He observed that brands which may have approximately the same attribute values (e.g., Coke and Pepsi, alternative physicians) sometimes have very different market shares. Consequently, he proposed to model the "overall utility" an individual received from a product as a sum of the "utility" estimated by a multiattribute model (developed from individual preferences) plus a brand-specific effect (one which is not individual-specific). He reported limited empirical results which favored his model over a logistic one similar to the logit model.

Finally, the approach taken by Urban (1975) presents both an interesting way of dealing with choice and a potentially important extension of the conceptual framework. In predicting choice or preference for a new brand on the basis of ideal point to product distance, one is effectively forecasting what might happen after customers have become familiar with the new brand and its image has become "stable." Urban defined this long-run market share for a new brand as the product of "the fraction of the target market who try the brand" and the "market share of purchases of the new brand given trial." Probability of trial was postulated to be an inverse function of ideal point to product distance defined only on the perceptual product space of persons who are merely aware of the new brand but have not used it. He also suggested that the equilibrium of a two-state Markov process might be used to estimate the market share of purchases of the new

brand after trial. The repeat purchase probability in the Markov model would be linked to distance defined now on the space of persons who have actually used the brand. And, the brand-switching probability in the Markov model would be estimated empirically from test market data. Urban thus proposed a multistage data collection process to obtain preference data after concept awareness and again after in-home usage had occurred. The value of multistage data collection in relation to its added complication and expense requires further investigation. Clearly such an approach is only useful for concept evaluation and not for generation because it will not be possible to familiarize respondents with all candidate concepts to obtain the respective switching probabilities.

In summary, treating choice as a stochastic variable relatable to a deterministic measure of preference by a separate modeling step appears to make good empirical sense. Either logit, Pessemier's model, or Srinivasan's brand-specific effects model seems useful. Confining a nonzero probability of choice to a subset of possible brands (which may vary with the individual) adds realism, as evoked set sizes tend to be small (Urban 1975). Because the marketer is ultimately interested in aggregate level prediction, modeling individuals as though they made but a single choice may simplify analysis without significant loss in predictive validity.

5. Evaluation and generation of product concepts. Within a situational submarket, evaluation of new product concepts generally has proceeded by positioning each alternative concept within the space of existing product/brands, calibrating choice models representative of each major market segment, and then determining the share of each segment's choices received by the new concept and noting from which existing product/brands these are drawn. Such shares can be weighted by a measure of each segment's purchasing power for products in the class (and by the relative frequency of occurrence of situations) and aggregated over segments and submarkets to predict overall market performance. If prespecified concepts are not available, this framework can be used to generate them through optimum-seeking procedures designed to search feasible regions of perceptual space for "promising" candidates. Search requires that managerial objectives be operationalized and that procedures be available for quickly evaluating relatively large numbers of alternatives.

Positioning a preexisting concept in the perceptual space can be accomplished by asking a sample of respondents to scale the concept description on the same attribute scales originally used to position existing products. Fairly straightforward transformations can be applied, in principle, for finding the concept's reduced space location, although the validity of such procedures can be questioned because techniques used to develop reduced space configurations are statistical processes. A new concept added to an existing space may not have the same location it would have had if it had

been included among the set of product alternatives used when the space was first determined (Huber and Reibstein 1977). Whether this observation is consequential is an empirical question, but it serves again to emphasize the caution which should be taken in working with reduced space configurations.

Market Segmentation. Evaluation within a situational submarket is facilitated by not having to work with many individual customer models. Rather, models of a few customer segments would prove useful. Several new approaches to the creation of segments appear promising. Pessemier (1975) suggested using dollar metric scaled preferences as the basis for forming "product preference market segments." Each segment would be composed of individuals who had common relative preferences for all existing products. Thus, in principle, there could be as many segments as there were distinct preference orderings among the existing products (i.e., all possible permutations of the product/brand alternatives). In practice, one would expect either that many possible orderings would not be empirically consequential or that the more reliable estimatable alternatives (most and least preferred rather than those in the middle) could be used to create a relatively small number of segments (Best 1978). Urban (1975) suggested defining segments on the basis of "similar" distances between the brands and the individual ideal points. Such a procedure might be expected to result in fewer segments than would be obtained by analyzing preference orderings because certain reversals in orderings may not be consequential in estimating ideal points or attribute saliences. Empirical evaluation of these alternative segmentation procedures for a firm's marketing purposes would be desirable.

Other methods for segmentation are possible. Segmentation on the basis of ideal point locations (Johnson 1971) or attribute saliences (Lehmann 1971) alone ignores the other aspect of the decision model and hence appears to offer a less desirable criterion than those mentioned before. When available data are insufficient to model at the individual level, more conventional segmentation variables based on usage rates or socioeconomic, demographic, or psychographic factors (Frank, Massy, and Wind 1972; Hauser and Urban 1977; Urban 1975) may be the only feasible criteria. The question of what basis to use in creating segments is an important one. Whatever basis is chosen, it should be strongly related to behavior under alternative usage conditions. Segmentation in terms of decisions made or the decision model derived from those decisions would thus appear desirable. Conventional segmentation criteria may still prove useful, however, as discriminating variables to help distinguish the segments and thus permit greater generalization of research findings.

Srinivasan and Shocker (1973*a*; Srinivasan, Shocker, and Weinstein 1973) developed a method of pooling the paired comparison preference

judgments of all individuals characterized as belonging to the same segment to estimate a single evaluation model for the segment. Pooling of ordinal preference judgments appears to be more defensible than simple averaging of ideal points and attribute salience estimates from individual models because the ordinal judgments are comparable in meaning across individuals whereas the estimates are ratio scaled and hence unique only up to multiplication by an arbitrary positive scalar. Consequently, comparison of such estimates across individuals may not be valid without an arbitrary rescaling in an attempt to facilitate interpersonal comparisons of utilities (Parker and Srinivasan 1976).

Evaluation of New Concepts. Each candidate concept can be evaluated in terms of the share of preferences (choices) it would receive from each segment (in each situational submarket). Lehmann (1971) suggested estimates be made of the expected quantity purchased in that submarket by the individuals composing each market segment. This quantity then could be apportioned among the existing brands (and the new concept) in accordance with relative preferences or choices. Shocker and Srinivasan (1974) in turn suggested converting units demand into dollar equivalents to facilitate economic analysis. If development, manufacturing, and marketing costs for each candidate concept as well as the costs of competing brands marketed by the firm were known, incremental profit contribution could be used as a criterion for evaluating a new alternative. They argued that for cases in which the firm has other brands already in the market, share of market for the new concept may not be the most worthwhile criterion. Rather, a firm might wish to look at incremental profits after considering cannibalization of the firm's present brands. Hauser and Urban (1977) indicated that what is being predicted is ultimate or long-range share under assumptions of familiarity with the new concept and stable competitive marketing activity. They would adjust aggregate share of choices for each alternative by measures of the eventual levels of awareness and recall that the firm's contemplated marketing effort is expected to achieve.

Alternative Methods for Generating New Concepts. Pessemier (1975) discussed a specific approach to generating new concepts. His STRATOP program required as input the units demand expected from each segment, the proposed selling price per unit, and the estimated cost (assumed constant) of a unit of each design attribute (original attribute). Heuristic methods were used to identify promising locations in reduced space. Noting that, in principle, a very large number of combinations of original attributes can map onto a single reduced space location, he made use of linear programming in STRATOP to find the "least cost" mix of original attributes corresponding to the "optimal" reduced space location for a new concept. His assumption, however, that unit costs for each design attribute could be

readily estimated and remained approximately constant was necessary for the use of linear programming, but otherwise questionable. Such costs may be difficult to estimate when attributes are psychological or otherwise nonactionable. Moreover, they may be decidedly nonlinear because the cost of each design attribute may be dependent on levels of the other attributes or on demand for the different product concepts themselves (i.e., cost-volume relations). Notwithstanding such criticisms, in light of the paucity of approaches which consider costs at all, the pragmatic question is whether their incorporation in the analysis, even crudely, is worthwhile.

A more serious problem with Pessemier's approach is his contention that there is some kind of structural relation between the perceptual space in terms of original attributes and its reduced space representation (obtained through a multivariate technique such as discriminant analysis). Though there may well be different cost means of creating any given reduced space location, the assumption of constant preference for the infinity of points in the original space (corresponding to a given reduced space location) is very questionable.

Analytic generation of new product concepts involves the formal process of seeking locations within technically and economically feasible regions of the perceptual space which optimize a managerially relevant objective. (None of the approaches provide guidance to the specification of such regions.) Shocker and Srinivasan (1974) posed this task as a nonlinear programming problem but indicated that assumptions necessary to make their model tractable would also make it unrealistic. They suggested instead two heuristic approaches: (1) modified "gradient search" (i.e., beginning at an arbitrary feasible location in the space and moving fixed steps in whichever direction improved the value of the objective until no further improvement was possible by a step increment in any direction—a local optimum), and (2) "search through coarse and fine grids" (i.e., dividing the perceptual space into an arbitrary number of gridlike regions and evaluating a test concept located at each grid centroid; those grids having centroids favorably evaluated by this procedure would themselves be divided into finer grids and the process repeated). Pessemier's STRATOP program utilized a gradient heuristic.

Recent independent work by Zufryden (1976) and by Albers and Brockhoff (1977) has led to mixed integer nonlinear programming models for positioning single new products. Both procedures assume individuals have finite ideal points and make single choices among products (they will choose whichever product is closer to their ideal points). These researchers differ in their willingness to consider idiosyncratic attribute saliences—Albers and Brockhoff (1977) ignore them but indicate their model can be reformulated to include them—and in the specifics of their solution procedures. Each approach seeks a single location (one new concept) which will have maximum sales. Zufryden uses linear approximations to each

customer's ellipsoidal acceptance zone (a consequence of allowing different attribute saliences) in order to apply available optimization techniques. Albers and Brockhoff, however, work directly with spheroidal zones. Zufryden does not report computational results. Times for Albers and Brockhoff procedures are large and appear to grow exponentially with numbers of customers in the sample (reaching more than one hour of CPU time on the PDP 10 for a problem with 200 customers). This fact suggests a need to work either with a small number of customer segments and/or to use heuristic approximation procedures for realistically sized problems.

By the approaches of Zufryden and Albers and Brockhoff, one cannot readily constrain the search to technically and economically feasible regions. This can be a serious limitation because many of the "optimal" new products identified may not be practical. Also, these approaches have only been formulated for objectives of maximum sales for the new brand rather than maximum incremental profits from an entire line (although Albers and Brockhoff claim to have a formulation which at least considers the costs of assessing various locations in the attribute space). The heuristic approaches, in contrast, have the flexibility to take into account technological feasibility considerations, different managerial objectives, and nonlinear costs. Among such heuristic procedures, modified "gradient search" appears less desirable than "search through coarse and fine grids" because of local optima, although this difficulty could be minimized by trying several starting points. (Albers and Brockhoff suggest such problems with gradient search by noting that Pessemier's STRATOP, which uses such heuristics, concentrates on neighborhoods of ideal or real products—a local optimum.) "Search through coarse and fine grids," however, may be computationally feasible only when the number of attributes is small.

Conclusions. In their present state of development, the approaches may be more appropriate for evaluating preexisting concepts or for product refinement (modifying an existing design to improve its market acceptance) than for concept generation. Whether perceptual spaces are always continuous, whether small changes in a product's spatial position will even be perceived, and whether decision models will continue to predict reasonably well over the entire domain of new product possibilities are some of the empirical questions which await additional research (Shocker, May, and Sudharshan 1978).

Implementation of the Different Approaches

Empirical applications of these approaches have begun to be reported. Stefflre looked at about 20 consumer markets, ranging from $200 million to $2 billion in volume, and tested about 150 new product descriptions (1971,

p. 59). He reported having fully developed six products which reached commercialization or test market and assisted in the development or introduction of an undisclosed additional number. Wind, Myers, Johnson, Pessemier, and Hustad et al. strongly hinted that clients have used their approaches and found them valuable. Ryans (1974) showed good predictive results for a nonlinear compensatory model in the context of predicting preferences for a durable good (blenders) at different prices. The linear compensatory model has been shown to provide good predictive results for adoption of a new men's shaving concept (Braun and Srinivasan 1975) and with preferences for breakfast cereals (Pekelman and Sen 1974). Green and Srinivasan (1978), in a recent review of applications of conjoint analysis, report an increasing number of additional applications. The complete Shocker and Srinivasan framework has been validated in the context of planning rural primary health care facilities (Parker and Srinivasan 1976). Urban (1975) summarized several empirical tests involving eight new product concepts for frequently purchased consumer nondurables. These tests investigated the predictive validity of model constructs and of the overall structure. He observed that such models provide managers with a better understanding of the perception, preference, and purchase structure of their markets, meet specific measurement needs, serve as a structure for interpretation of experimental results, and aid managers in channeling their creative effort to develop successful new product designs. Finally, Hauser and Urban (1977) reported good results in applying their procedures to the design of a health maintenance organization and a new deodorant. They also commented on applications of their approach to the design of a Masters degree program at MIT, the design of financial service packages, and the positioning of several new frequently purchased consumer products.

Conclusions

This review suggests that much has been accomplished, much promise shown. But as is usual in the early development of a field, much work remains to be done. Many of the approaches identified are as yet unfulfilled promises, the researchers having been content to suggest procedures on the basis of personal judgments without being concerned with issues of measure reliability and validity or the validity of the conceptual frameworks themselves. It is particularly disturbing to find that the problems of identifying relevant usage situations, the composition of product-markets, and determinant attributes from customer perspectives are so frequently ignored.

The concept of the usage situation needs greater refinement to embrace

products which provide multiple benefits simultaneously (e.g., for many customers a general purpose cleanser may or may not be highly competitive with special purpose products for each special use) or continue their benefits over protracted time periods (e.g., an automobile may be used in a large number of "situations" which may not be fully anticipated at time of purchase) or which can be used in alternative ways. It is necessary to compare and test different methods for creating situational typologies. The relation between multistage decision making and the definition of product-market boundaries for situational submarkets should be examined (Day, Shocker, and Srivastava 1978; Srinivasan 1978; Srivastava, Shocker, and Day 1978). The authors have hypothesized, for example, that intended usage determines criteria which serve to restrict an individual's consideration set and that individuals make final choices only from brands in that restricted set. This implies that individuals go through at least a two-phase decision process, employing different criteria at each phase and possibly different decision rules. A stronger theoretical and empirical basis for such conjectures would be very useful to manufacturers wishing to learn what they must do to have their brands considered for purchase under different situations. Further, methods for aggregating analyses from different situational submarkets into a total market analysis have not been explicitly detailed.

Additional research is needed to identify the circumstances under which perceptual heterogeneity may be serious. One factor seeming to affect homogeneity is the extent to which individuals are familiar or otherwise have experience with the existing product alternatives (Silk and Urban 1976; Srivastava, Shocker, and Day 1978). Thus the assumption of homogeneity may only be reasonable for established markets (or industrial settings) where such familiarity can be reasonably expected. In other instances it may be desirable to impose additional screening criteria in selecting samples of respondents or to make use of descriptions of existing alternatives as well as new concepts (and thus better control the information on which respondents base their responses). Besides a greater likelihood of perceptual homogeneity, the use of descriptions offers the benefit of more reliable calibration of the individual decision models (e.g., greater variation in underlying attribute levels than may be offered by existing brands). However, use of descriptions as stimuli raises the additional question of just what are the better ways to create valid representations of real objects (Green and Srinivasan 1978). Research in this area may also shed light on the question of actionability which, though critical to concept generation, also directly affects the task of creating real objects and marketing actions from descriptions (Green 1975; Shocker and Srinivasan 1974).

Much empirical research is needed to determine which methods are more appropriate for each stage in the conceptual framework and under what conditions. The process of obtaining preference judgments for

hypothetical descriptions and the subsequent estimation of decision model parameters from these data require closer scrutiny (Green and Srinivasan 1978). The diagnostic value of such parameter estimates needs to be established in a behaviorally valid (Shocker and Zaltman 1977) way (e.g., one which demonstrates that products at or near ideal point locations are really purchased more frequently than those farther away or that improving a brand's perceived level on an "important" attribute actually leads to increased sales for it). Customer decision models offer intriguing possibilities for defining market segments more in accord with the conceptual basis for segmentation, i.e., that members of each segment will react similarly to the marketing efforts of the firm (Frank, Massy, and Wind 1972). Specific criteria for such segmentation have been suggested in connection with some of the different approaches reviewed but their behavioral validity has yet to be established.

The usefulness of these frameworks under more dynamic conditions is just beginning to be examined (Silk and Urban 1976). A perceptual and preference framework nominally offers a static description of current market conditions. No explanation is provided for how perceptual and decision structures are created through learning and how they are modified by further experience and environmental change. For example, Pessemier (1975) along with others (Green and Carmone 1970) suggested that these frameworks could be regarded as market simulations in which various marketing and product strategies might be evaluated. The validity of this idea requires answers to such questions as (Shocker, May, and Sudharshan 1978): How stable are the perceptual space and individual decision models when relatively large numbers of changes in market structure are to be evaluated, e.g., when the firm's moves, competitor's reactions, and the firm's countermoves are being modeled (Rao 1977) or when a large number of new concepts are to be added or existing products removed from the market simultaneously (a point Pessemier seems not to question)?[5] If a number of new products are added or existing ones "improved" it may be reasonable to expect the total size of market to grow either as a function of increased convenience resulting from greater availability of purchasable alternatives or of increased overall product quality (Huber and Reibstein 1977). When a product is first introduced, its benefits and costs will not be immediately understood or appreciated by all market participants. It seems necessary to incorporate the effects of initial marketing expenditures in increasing the level of awareness and availability of the product. Can a new concept be modeled in the perceptual space as a *region of uncertainty* (Ahtola 1975) whose domain may be increasingly better approximated by a point as a result of experience with a brand and/or awareness and knowledge created by promotional activities? The concept of preference inertia introduced by Neslin (1976) might be helpful in relating steady-state

market shares to transient market behavior. An assumption implicit in the approaches reviewed is that concepts with higher long-term or steady-state potential would also perform better in the short term. But performance depends in part on the level and quality of marketing effort the firm undertakes. This observation suggests difficulties in validating the framework, because whatever results are observed are ambiguous. If an "optimal" new product concept is not immediately successful, is it because the criteria used to discover it are invalid or because it was executed improperly (Brown et al. 1968)? And if the product does not meet the firm's short-term profit objectives, will the firm stick with it long enough to assess its ultimate potential?

Finally, it must be emphasized that the strength of these approaches lies in the fact that they are customer-based. Customer inputs at an early stage of managerial planning can increase management's identification with its customers' desires and behaviors. Moreover, such inputs at an early stage can preclude a management's tendency to resolve uncertainties "conveniently" by failing to conduct research at all or by dictating its "results" in advance, pursuing actions merely because the firm finds them easy to accomplish, or developing marketing strategy influenced unduly by the status or "force of personality" of certain key executives or by their "pet projects." The conceptualizations of the perceptual product space, ideal points, and distance-related models of preference lend themselves to geometric representation which can aid management in visualizing the marketing environment and creating promising strategy (Brown et al. 1968; Johnson 1971; Srivastava, Shocker, and Day 1978). Such a framework also affords the analyst a systematic way of tracking parameter values over time and relating such changes to developments in the marketplace. It presents an enormous opportunity for a firm to *learn*, to acquire knowledge systematically by modeling the effects of its own and competitors' actions and those of the external sociopolitical environment on behaviors in the market. The result should be an improved management information system and an ability to manage the enterprise.

The frameworks reviewed represent unfulfilled promises. Whether they can accomplish all that is hoped for is a researchable question. But even if the approaches could be totally validated, will managements use them? Too much market knowledge may be threatening because it can restrict managerial prerogatives and interfere with relationships of power and expertise which have been built up over many years. Implementation calls for new perspectives and the understanding or acceptance of certain analytical approaches with which few managers may be familiar. Stefflre and others have commented very critically on the organizational problems which impede new-product planning efforts (Day 1968; Stefflre 1971). The critical question is what relative advantage these newer approaches have over current, more conventional approaches to product and marketing

planning. Demonstration of a compelling competitive advantage to the firms which use them would be the most important factor favoring their adoption.

Notes

1. Real product-markets often consist of brands which are very similar to each other. If a product is "successful," one may expect other products/brands to be positioned similarly. Many manufacturers consciously introduce "me too" brands. Consequently, for purposes of improving the precision of results at later stages of the framework, it will often be advisable to create fictitious products (descriptions) which produce more variation along the attributes.

2. An earlier version appeared in an article by Pessemier and Root (1973).

3. Myers and Shocker (1978) also argued that because ideal products need not have attribute levels which lie within the same range as existing ones, the use of the same bipolar scales for scaling both may result in respondents treating the scales differently in the two instances.

4. Briefly, if two substantially different products, say automobiles (A) and public transit (B), have market shares .75 and .25, respectively, and another mode of public transit (C) very similar to B is introduced, the logit model leads to market shares of .6, .2, .2 (the total share for transit has risen). This outcome is intuitively unappealing. Because the two transit modes B and C are very similar, the introduction of C should draw market share largely from B rather than A.

5. Ryans (1974) found a relatively high level of stability when a single new concept was introduced into an existing product-market.

References

Ahtola, Olli (1975), "The Vector Model of Preferences: An Alternative to the Fishbein Model," *Journal of Marketing Research,* 12(February), 52–9.

Albers, Sonke, and Klaus Brockhoff (1977), "A Procedure for New Product Positioning in an Attribute Space," *European Journal of Operational Research,* Volume 1, 230–8.

Alpert, Mark I. (1971), "Definition of Determinant Attributes: A Comparison of Methods," *Journal of Marketing Research,* 8 (May), 184–91.

Alpert, M., John F. Betak, and Linda L. Golden (1978), "Data Gathering

Issues in Conjoint Measurement," Working Paper, Graduate School of Business, The University of Texas.

Beckwith, Neil E., and Donald R. Lehmann (1975), "The Importance of Halo Effects in Multi-Attribute Attitude Models," *Journal of Marketing Research,* 12 (August), 265-75.

Belk, Russel W. (1975), "Situational Variables and Consumer Behavior," *Journal of Consumer Research,* 2 (December), 157-64.

Benson, Purnell H. (1966), "Analysis of Irregular Two Dimensional Distributions of Consumer Buying Choice," *Journal of Marketing Research,* 3 (August), 278-88.

——— (1965), "Fitting and Analyzing Distribution Curves of Consumer Choices," *Journal of Advertising Research,* 5 (March), 28-34.

Berkowitz, Eric N., James L. Ginter, and W. Wayne Talarzyk (1975), "Incorporating Specific Use Situations Data in a Multi-Attribute Model for the Prediction of Individual Brand Choice," Working Paper, College of Administrative Sciences, Ohio State University.

Best, Roger J. (1978), "Validity and Reliability of Criterion-Based Preferences," *Journal of Marketing Research,* 5 (February), 154-60.

Bettman, James R. (1971), "The Structure of Consumer Choice Processes," *Journal of Marketing Research,* 8 (November), 465-71.

Braun, Michael A., and V. Srinivasan (1975), "Amount of Information as a Determinant of Consumer Behavior Toward New Products," *1975 Combined Proceedings,* American Marketing Association, 373-8.

Brown, Milton P., Richard N. Cardozo, Scott M. Cunningham, Walter J. Salmon, and Ralph G.M. Sultan (1968), "Maxwell House Division," pp. 439-66 in *Problems in Marketing.* New York: McGraw-Hill Book Company.

Cattin, Phillipe, and Dick R. Wittink (1976), "A Monte Carlo Study of Metric and Non-Metric Estimation Methods for Multi-attribute Models," Research Paper No. 341, Graduate School of Business, Stanford University (November).

Dawes, Robyn, and Bernard Corrigan (1974), "Linear Models in Decision-Making," *Psychological Bulletin,* 81 (February) 95-106.

Day, George S., Terry Deutscher, and Adrian B. Ryans (1976), "Data Quality, Level of Aggregation, and Nonmetric Multidimensional Scaling Solutions," *Journal of Marketing Research,* 13 (February), 92-7.

Day, G.S., Allan D. Shocker, and Raj Srivastava (1978), "Identifying Competitive Product-Markets: A Review of Customer-Oriented Approaches," Working Paper, Graduate School of Business, University of Pittsburgh (June).

Day, Ralph (1968), "Preference Tests and the Management of Product Features," *Journal of Marketing,* 32 (July), 24-9.

Deutscher, Terry, and Adrian B. Ryans (1975), "Joint Space Mapping as a

Means of Understanding Changes in Brand Preference," Working Paper 75-27, College of Administrative Sciences, Ohio State University (August).

Fishbein, Martin, and Icek Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research,* Reading, Massachusetts: Addison-Wesley.

Frank, Ronald E., William F. Massy, and Yoram Wind (1972), *Market Segmentation.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Green, Paul E. (1975), "Marketing Applications of MDS: Assessment and Outlook," *Journal of Marketing* 39 (January), 24–31.

Green, P.E., and Frank Carmone (1970), *Multidimensional Scaling and Related Techniques in Marketing Analysis.* Boston: Allyn and Bacon.

Green, P.E., and Michael T. Devita (1975), "The Robustness of Linear Models under Correlated Attribute Conditions," *1975 Combined Proceedings,* American Marketing Association.

Green, P.E., and Vithala R. Rao (1972), *Applied Multidimensional Scaling.* New York: Holt Rinehart and Winston.

Green, P.E., and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research,* 5 (September).

Green, P.E., and Yoram Wind (1973), *Multi-Attribute Decisions in Marketing,* Hinsdale, Illinois: The Dryden Press.

Green, P.E., Yoram Wind, and Arun K. Jain (1972), "Benefit Bundle Analysis," *Journal of Advertising Research,* 12 (April), 31-6.

Haley, Russell I. (1968), "Benefit Segmentation: A Decision-Oriented Research Tool," *Journal of Marketing,* 32 (July), 30–5.

Hansen, Flemming (1976), "Psychological Theories of Consumer Choice," *Journal of Consumer Research,* 3 (December), 117–42.

Hauser, John R., and Frank S. Koppelman (1979), "Effective Marketing Research: An Empirical Comparison of Techniques to Model Consumers' Perceptions and Preferences," in *Analytic Approaches to Product and Marketing Planning,* Allan D. Shocker, ed. Cambridge, Massachuetts: Marketing Science Institute.

Hauser, J.R., and Steven M. Shugan (1977), "Efficient Measurement of Consumer Preference Functions: A General Theory of Intensity of Preference," Working Paper 602-001, Graduate School of Management, Northwestern University (July).

Hauser, J.R, and Glen L. Urban (1976), "Direct Assessment of Consumer Utility Functions: Von Neumann-Morgenstern Utility Theory Applied to Marketing," Working Paper 76-17, Graduate School of Management, Northwestern University (March).

——— (1977), "A Normative Methodology for Modeling Consumer Response to Innovation," *Operations Research,* 25 (July–August), 579–619.

Howard, John A., and Jagdish Sheth (1969), *The Theory of Buyer Behavior.* New York: John Wiley & Sons, Inc., 191–228.

Huber, George P. (1974), "Multi-Attribute Utility Models: A Review of Field and Field-Like Studies," *Management Science,* 20 (June), 1393–1402.

Huber, Joel (1975), "Predicting Preference on Experimental Bundles of Attributes: A Comparison of Models," *Journal of Marketing Research,* 12 (August), 290–7.

Huber, J., and David J. Reibstein (1977), "Attitude Measures and Choice Frequency: Some Pitfalls to Be Avoided," Paper presented at AMA Attitude Research Conference, Las Vegas (March).

Hustad, Thomas P., Charles S. Mayer, and Thomas W. Whipple (1975), "Consideration of Context Differences in Product Evaluation and Market Segmentation," *Journal of the Academy of Marketing Science,* 3 (Winter), 34–47.

Jain, Arun, Franklin Acito, Naresh Malhotra, and Vijay Mahajan (1978), "A Comparison of Predictive Validity of Alternate Methods for Estimating Parameters in Preference Models," Working Paper, School of Management, State University of New York (March).

Johnson, Richard M. (1970), "Multiple Discriminant Analysis: Applications to Marketing Research." Chicago: Market Facts, Inc. (January).

——— (1971), "Market Segmentation: A Strategic Management Tool," *Journal of Marketing Research,* 8 (February), 13–18.

——— (1974), "Tradeoff Analysis of Consumer Values," *Journal of Marketing Research,* 11 (May), 121–7.

Kinnear, Thomas C., and James R. Taylor (1973), "The Effect of Ecological Concern on Brand Perceptions," *Journal of Marketing Research,* 10 (May), 191–7.

Klahr, David (1970), "A Study of Consumer Cognitive Structure for Cigarette Brands," *Journal of Business,* 43 (April), 190–204.

Kuehn, Alfred A., and Ralph L. Day (1962), "Strategy of Product Quality," *Harvard Business Review,* 40 (November–December), 100–10.

Lehmann, Donald R. (1971), "Evaluating Marketing Strategy in a Multiple Brand Market," *Journal of Business Administration,* 3 (Fall) 15–26.

Lutz, Richard J., and Pradeep Kakkar (1975), "The Psychological Situation as a Determinant of Consumer Behavior," in *Advances in Consumer Research,* Mary Jane Schlinger, ed. Chicago: Association for Consumer Research, 439–53.

McFadden, Daniel (1970), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics,* P. Zarembka, ed. New York: Academic Press, 105–42.

Miller, Kenneth E., and James L. Ginter (1977), "An Investigation of Situational Variance of Brand Choice Behavior and Attitude," Working Paper, School of Business, University of Santa Clara (March).

Montgomery, David B., and Glen L. Urban (1969), *Management Science in Marketing*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 53–84.

Montgomery, D.B., Dick R. Wittink, and Thomas Glaze (1978), "A Predictive Test of Individual Level Concept Evaluation and Tradeoff Analysis," Research Paper No. 415, Graduate School of Business, Stanford University.

Morgan, Nigel, and Janet Purnell (1969), "Isolating Openings for New Products in a Multidimensional Space," *Journal of the Market Research Society,* 11 (July), 245–66.

Myers, James H. (1976), "Benefit Structure Analysis: A New Tool for Product Planning," *Journal of Marketing,* 40 (October), 23–32.

——— (1970), "Finding Determinant Buying Attitudes," *Journal of Advertising Research,* 1 (December), 9–12.

Myers, J.H., and Richard F. Chay (1977), "Direct vs. Derived Estimates of Ideal Product Characteristics," Working Paper, School of Business Administration, University of Southern California (February).

Myers, J.H., and Allan D. Shocker (1978), "Toward a Taxonomy of Product Attributes," Working Paper, 320, Graduate School of Business, University of Pittsburgh (October).

Nakanishi, Masao, Lee G. Cooper, and Harold H. Kassarjian (1974), "Voting for a Political Candidate Under Conditions of Minimal Information," *Journal of Consumer Research,* 1 (September), 36–43.

Neslin, Scott A. (1976), "Analyzing Consumer Response to Innovation: The Concept of Preference Inertia," Working Paper 877–76, Sloan School of Management, M.I.T. (October).

Parker, Barnett R., and V. Srinivasan (1976), "A Consumer Preference Approach to the Planning of Rural Primary Health Care Facilities," *Operations Research,* 24 (September-October), 991–1025.

Perkelman, Dov, and Subrata Sen (1974), "Mathematical Programming for the Determination of Attribute Weights," *Management Science,* 20 (April), 1217–29.

Pessemier, Edgar A. (1975), "Market Structure Analysis of New Product and Market Opportunities," *Journal of Contemporary Business* (Spring), 35–67.

Pessemier, E.A., Philip Burger, Richard Teach, and Douglas Tigert (1971), "Using Laboratory Preference Scales to Predict Consumer Brand Preferences," *Management Science,* 17 (February), 371–85.

Pessemier, E.A., and Paul Root (1973), "The Dimensions of New Product Planning," *Journal of Marketing,* 37 (January), 10–18.

Punj, Girish, and Richard Staelin (1978), "The Choice Process for Graduate Business Schools," *Journal of Marketing Research,* 15 (November), 588–98.

Rao, Ram C. (1977), "Rivalistic Analysis for Evaluating New Product Concepts," in *Production Management with Interfaces to Marketing and Oligopolistic Competition,* unpublished doctoral dissertation,

Graduate School of Industrial Administration, Carnegie-Mellon University (May).

Rao, Vithala, and Geoffrey N. Soutar (1975), "Subjective Evaluations for Product Design Decisions," *Decision Sciences,* 6 (January), 120–34.

Ritchie, J. Brent (1974), "An Exploratory Analysis of the Nature and Extent of Individual Differences in Perception," *Journal of Marketing Research,* 11 (February), 41–9.

Ryan, Michael J., and Edward H. Bonfield (1975), "The Fishbein Extended Model and Consumer Behavior," *Journal of Consumer Research,* 2 (September), 118–36.

Ryans, Adrian B. (1974), "Estimating Customer Preferences for a New Durable Brand in an Established Product Class," *Journal of Marketing Research,* 11 (November), 431–43.

Shocker, Allan D., Dennis Gensch, and Leonard Simon (1969), "Toward the Improvement of New Product Search and Screening," *Proceedings of the Fall Conference,* American Marketing Association, 168–75.

Shocker, A.D., Jerrold May, and D. Sudharshan (1978), "Design of New Product Concepts: A Comparison of Optimizing Approaches," Paper presented at Joint National ORSA/TIMS Conference, Los Angeles (November).

Shocker, A.D., and V. Srinivasan (1974), "A Consumer-Based Methodology for the Identification of New Product Ideas," *Management Science,* 20 (February), 921–38.

Shocker, A.D., and Gerald Zaltman (1977), "Validity Importance in Consumer Research: Some Pragmatic Issues" in *Advances in Consumer Research,* William Perrault, ed. Atlanta: Association for Consumer Research, 405–8.

Silk, Alvin J., and Glen L. Urban (1976), "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," Working Paper 834–76, Alfred P. Sloan School of Management, M.I.T. (January).

Srinivasan, V. (1975), " A General Procedure for Estimating Consumer Preference Distributions," *Journal of Marketing Research,* 12 (November), 337–89.

——— (1978), "A Model and Estimation Procedure for Multi-Stage Decision Processes," Working Paper, Graduate School of Business, Stanford University (May).

——— (1976), "Network Models for Estimating Brand Images," Research Paper No. 344, Graduate School of Business, Stanford University (December).

Srinivasan, V., and Allan D. Shocker (1973*a*), "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments," *Psychometrika,* 38 (December), 473–93.

——— (1973*b*), "Linear Programming Techniques for Multi-Dimensional Analysis of Preferences," *Psychometrika,* 38 (September), 337–69.

Srinivasan, V., Allan D. Shocker, and Alan G. Weinstein (1973), "Measurement of a Composite Criterion of Managerial Success," *Organizational Behavior and Human Performance,* 9 (February), 147-67.

Srivastava, Rajendra, Allan D. Shocker, and George Day (1978), "An Exploratory Study of Situational Effects on Product-Market Definition," in *Advances in Consumer Research,* H. Keith Hunt, ed. Atlanta: Association for Consumer Research.

Stefflre, Volney J. (1971), *New Products and New Enterprises: A Report of an Experiment in Applied Social Science.* Irvine, California: University of California, (March).

——— (1972), "Some Applications of Multi-Dimensional Scaling to Social Science Problems, in *Multi-dimensional Scaling: Theory and Applications in the Behavioral Sciences,* Roger N. Shepard, A.K. Romney, and Sarah B. Nerlove, eds. New York: Seminar Press.

Urban, Glen L. (1975), "PERCEPTOR: A Model for Product Positioning," *Management Science,* 21 (April), 858-71.

Wildt, Albert R., and Albert V. Bruno (1974), "The Prediction of Preference for Capital Equipment Using Linear Attitude Models," *Journal of Marketing Research,* 11 (May) 203-5.

Wilkie, William L., and Edgar A Pessemier (1973), "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Research,* 10 (November), 428-41.

Wilkie, W.L., and Rolf P. Weinreich (1973), "Effects of the Number and Type of Attributes Included in an Attitude Model: More Is Not Better," Institute Paper No. 385, Krannart Graduate School of Administration, Purdue University.

Wind, Yoram (1973), "A New Procedure for Concept Evaluation," *Journal of Marketing,* 37 (October), 2-11.

Wright, Peter L. (1975), "Consumer Choice Strategies: Simplifying vs. Optimizing," *Journal of Marketing Research,* 12 (February), 60-7.

——— (1973), "Use of Consumer Judgment Models in Promotion Planning," *Journal of Marketing,* 37 (October), 27-33.

Wright, P.L., and Frederick Barbour (1977), "Phased Decision Strategies: Sequels to an Initial Screening," in *Multiple Criteria Decision Making: TIMS Studies in the Management Sciences,* Vol. 6, Martin K. Starr and M. Zeleny, eds. (available as a special issue of *Management Science*). Amsterdam: North-Holland, 91-109.

Zufryden, Fred S. (1976), "ZIPMAP—A Zero-One Integer Programming Model for Market Segmentation and Product Positioning," Working Paper, Graduate School of Business Administration, University of Southern California.

3 New Way to Measure Consumers' Judgments

Paul E. Green and
Yoram Wind

Taking a jet plane for a business appointment in Paris? Which of the two flights described below would you choose?

> A B-707 flown by British Airways that will depart within two hours of the time you would like to leave and that is often late in arriving in Paris. The plane will make two intermediate stops, and it is anticipated that it will be 50% full. Flight attendants are "warm and friendly" and you would have a choice of two movies for entertainment.
>
> A B-747 flown by TWA that will depart within four hours of the time you would like to leave and that is almost never late in arriving in Paris. The flight is nonstop, and it is anticipated that the plane will be 90% full. Flight attendants are "cold and curt" and only magazines are provided for entertainment.

Are you looking for replacement tires for your two-year old car? Suppose you want radial tires and have the following three options to choose from:

> Goodyear's, with a tread life of 30,000 miles at a price of $40 per tire; the store is a 10-minute drive from your home.
>
> Firestone's, with a tread life of 50,000 miles at a price of $85 per tire; the store is a 20-minute drive from your home.
>
> Or Sear's, with a tread life of 40,000 miles at a price of $55 per tire; the store is located about 10 minutes from your home.

How would you rank these alternatives in order of preference:

Both of these problems have a common structure that companies and their marketing managers frequently encounter in trying to figure out what

---

a consumer really wants in a product or service. First, the characteristics of the alternatives that the consumer must choose from fall along more than a single dimension—they are multiattribute. Second, the consumer must make an overall judgment about the relative value of those characteristics, or attributes; in short, he must order them according to some criterion. But doing this requires complex trade-offs, since it is likely that no alternative is clearly better than another on every dimension of interest.

In recent years, researchers have developed a new measurement technique from the fields of mathematical psychology and psychometrics that can aid the marketing manager in sorting out the relative importance of a product's multidimensional attributes.[1] This technique, called conjoint measurement, starts with the consumer's overall or global judgments about a set of complex alternatives. It then performs the rather remarkable job of decomposing his or her original evaluations into separate and compatible utility scales by which the original global judgments (or others involving new combinations of attributes) can be reconstituted.[2]

Being able to separate overall judgments into psychological components in this manner can provide a manager information about the relative importance of various attributes of a product. It can also provide information about the value of various levels of a single attribute. (For example, if price is the attribute under consideration, conjoint measurement can give the manager a good idea of how sensitive consumers would be to a price change from a level of, say, 85¢ to one of 75¢ or one of 95¢.) Indeed, some models can even estimate the psychological trade-offs consumers make when they evaluate several attributes together.

The advantages of this type of knowledge to the planning of marketing strategy are significant. The knowledge can be useful in modifying current products or services and in designing new ones for selected buying publics.

In this article, we first show how conjoint measurement works from a numerical standpoint. We then discuss its application to a variety of marketing problems, and we demonstrate its use in strategic marketing simulations. Appendix 3A provides a brief description of how other research tools for measuring consumer judgments work, and how they relate to conjoint measurement.

How Conjoint Measurement Works

In order to see how to apply conjoint measurement, suppose a company were interested in marketing a new spot remover for carpets and upholstery. The technical staff has developed a new product that is designed to handle tough, stubborn spots. Management interest centers on five attributes or factors that it expects will influence consumer preference: an applicator-

type package design, brand name, price, a *Good Housekeeping* seal of endorsement, and a money-back guarantee.

Three package designs are under consideration and appear in the upper portion of figure 3-1. There are three brand names under consideration: *K2R, Glory,* and *Bissel.* Of the three brand names used in the study, two are competitors' brand names already on the market, whereas one is the company's present brand name choice for its new product. Three alternative prices being considered are $1.19, $1.39, and $1.59. Since there are three alternatives for each of these factors, they are called three-level factors. The *Good Housekeeping* seal and money-back guarantee are two-level factors, since each is either present or not. Consequently, a total of $3 \times 3 \times 3 \times 2 \times 2 = 108$ alternatives would have to be tested if the researcher were to array all possible combinations of the five attributes.

Clearly, the cost of administering a consumer evaluation study of this magnitude—not to mention the respondents' confusion and fatigue—would be prohibitive. As an alternative, however, the researcher can take advantage of a special experimental design, called an *orthogonal array,* in which the test combinations are selected so that the independent contributions of all five factors are balanced.[3] In this way each factor's weight is kept separate and is not confused with those of the other factors.

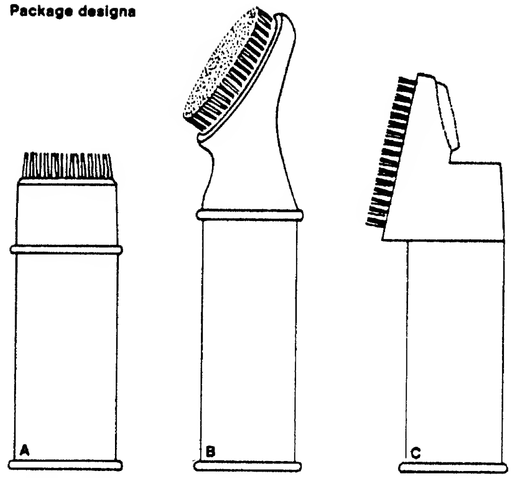
The lower portion of figure 3-1 shows an orthogonal array that involves only 18 of the 108 possible combinations that the company wishes to test in this case. For the test the researcher makes up 18 cards. On each card appears an artist's sketch of the package design, A, B, or C, and verbal details regarding each of the other four factors: brand name, price, *Good Housekeeping* seal (or not), and money-back guarantee (or not). After describing the new product's functions and special features, he shows the respondents each of the 18 cards (see figure 3-1 for the master design), and asks them to rank the cards in order of their likelihood of purchase.

The last column of figure 3-1 shows one respondent's actual ranking of the 18 cards; rank number 1 denotes her highest evaluated concept. Note particularly that only *ranked* data need be obtained and, furthermore, that only 18 (out of 108) combinations are evaluated.

Computing the Utilities

Computation of the utility scales of each attribute, which determine how influential each is in the consumers' evaluations, is carried out by various computer programs.[4] The ranked data of a single respondent (or the composite ranks of a group of respondents) are entered in the program. The computer then searches for a set of scale values for each factor in the experimental design. The scale values for each level of each factor are chosen so

Package designs



Orthogonal array

	Package design	Brand name	Price	Good Housekeeping seal?	Money-back guarantee?	Respondent's evaluation (rank number)
1	A	K2R	$1.19	No	No	13
2	A	Glory	1.39	No	Yes	11
3	A	Bissell	1.59	Yes	No	17
4	B	K2R	1.39	Yes	Yes	2
5	B	Glory	1.59	No	No	14
6	B	Bissell	1.19	No	No	3
7	C	K2R	1.59	No	Yes	12
8	C	Glory	1.19	Yes	No	7
9	C	Bissell	1.39	No	No	9
10	A	K2R	1.59	Yes	No	18
11	A	Glory	1.19	No	Yes	8
12	A	Bissell	1.39	No	No	15
13	B	K2R	1.19	No	No	4
14	B	Glory	1.39	Yes	No	6
15	B	Bissell	1.59	No	Yes	5
16	C	K2R	1.39	No	No	10
17	C	Glory	1.59	No	No	16
18	C	Bissell	1.19	Yes	Yes	1*

*Highest ranked

Figure 3-1. Experimental Design for Evaluation of a Carpet Cleaner

that when they are added together the *total* utility of each combination will correspond to the original ranks as closely as possible.

Notice that two problems are involved here. First, as mentioned previously, the experimental design of figure 3-1 shows only 18 of 108 combinations. Second, only rank-order data are supplied to the algorithms. This means that the data themselves do not determine how much more influential one attribute is than another in the consumers' choices. However, despite these limitations, the algorithms are able to find a *numerical* representation of the utilities, thus providing an indication of each factor's relative importance.

In general, more accurate solutions are obtained as the number of combinations being evaluated increases. Still, in the present case, with only 18 ranking-type judgments, the technique works well. Figure 3-2 shows the computer results.

As can be observed in figure 3-2, the technique obtains a utility function for each level of each factor. For example, to find the utility for the first combination in figure 3-1, we can read off the utilities of each factor level in the five charts of figure 3-2: U(A) = 0.1; U(K2R) = 0.3; U($1.19) = 1.0; U(No) = 0.2; U(No) = 0.2. Therefore the total utility is 1.8, the sum of the five separate utilities, for the first combination. Note that this combination was ranked only thirteenth by the respondent in figure 3-1.

On the other hand, the utility of combination 18 is 3.1 (0.6 + 0.5 + 1.0 + 0.3 + 0.7), which is the respondent's highest evaluation of all 18 combinations listed.

However, as can be easily seen from figure 3-2, if combination 18 is modified to include package Design B (in place of C), its utility is even higher. As a matter of fact, it then represents the highest possible utility, even though this specific combination did not appear among the original 18.

Importance of Attributes

By focusing attention on only the package design, the company's marketing researchers can see from figure 3-2 that Design B displays highest utility. Moreover, all utility scales are expressed in a common unit (although their zero points are arbitrary). This means that we can compare utility ranges from factor to factor so as to get some idea of their relative importance.

In the case of the spot remover, as shown in figure 3-2, the utility ranges are:

Package design (1.0 − 0.1 = 0.9)

Brand name (0.5 − 0.2 = 0.3)

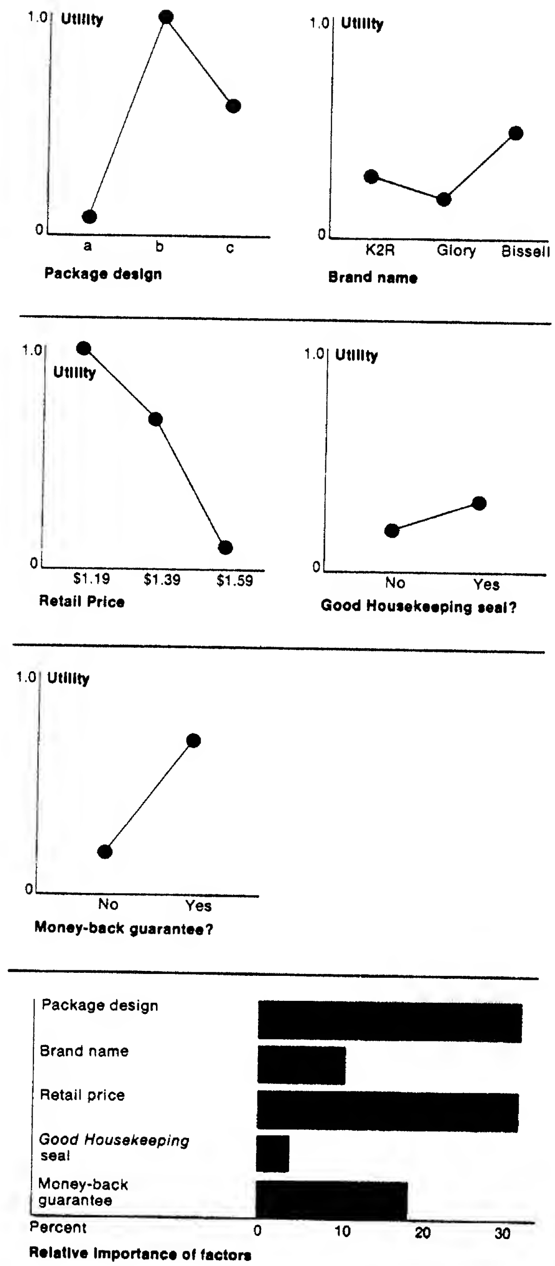


Figure 3–2. Results of Computer Analysis of Experimental Data of Figure 3–1

Price (1.0 − 0.1 = 0.9)

Good Housekeeping seal (0.3 − 0.2 = 0.1)

Money-back guarantee (0.7 − 0.2 = 0.5)

How important is each attribute in relation to the others? The lower portion of figure 3–2 shows the relative size of the utility ranges expressed in histogram form. As noted, package design and price are the most important factors, and together they account for about two thirds of the total range in utility.

It should be mentioned that the relative importance of a factor depends on the levels that are included in the design. For example, had price ranged from $1.19 to a high of $1.89, its relative importance could easily exceed that for package design. Still, as a crude indication of what factors to concentrate on, factor importance calculations provide a useful by-product of the main analysis regardless of such limitations.

Managerial Implications

From a marketing management point of view the critical question is how these results can be used in the design of a product/marketing strategy for the spot remover. Examination of figure 3–2 suggests a number of points for discussion:

> Excluding brand name, the most desirable offering would be the one based on package Design B with a money-back guarantee, a *Good Housekeeping* seal, and a retail price of $1.19.
>
> The utility of a product with a price of $1.39 would be 0.3 less than one with a price of $1.19. A money-back guarantee which involves an increment of 0.5 in utility would more than offset the effect of the higher price.
>
> The use of a *Good Housekeeping* seal of approval is associated with a minor increase in utility. Hence including it in the company's product will add little to the attractiveness of the spot remover's overall offering.
>
> The utility of the three brand names provides the company with a quantitative measure of the value of its own brand name as well as the brand names of its competitors.

Other questions can be answered as well by comparing various composites made up from the utilities shown in figure 3–2.

The Air Carrier Study

What about the two Paris flights you had to choose between? In that study, the sponsor was primarily interested in how air travelers evaluated the B-707 versus the B-747 in transatlantic travel, and whether relative value differed by length of flight and type of traveler—business versus vacation travelers. In this study all the respondents had flown across the Atlantic at least once during the preceding months.

Figure 3-3 shows one of the findings of the study for air travelers (business and vacation) flying to Paris. Without delving into details it is quite apparent that the utility difference between the B-707 and the B-747 is very small. Rather, the main factors are departure time, punctuality of arrival, number of stops, and the attitudes of flight attendants.

The importance of type of aircraft did increase slightly with length of flight and for business-oriented travelers versus vacationers. Still, its importance to overall utility was never greater than 10%. It became abundantly clear that extensive replacement of older aircraft like the B-707 would not result in major shifts in consumer demand. On the contrary, money might better be spent on improving the scheduling aspects of flights and the attitudes and demeanor of flight personnel.

The air carrier study involved the preparation of some 27 different flight profiles (only two of which appear at the beginning of the article). Respondents simply rated each flight description in terms of its desirability on a seven-point scale. Only the order properties of the ratings were used in the computer run that resulted in the utility scales appearing in figure 3-3.

The Replacement Tire Study

The conjoint measurement exercise in the replacement tire study was part of a larger study designed to pretest several television commercials for the sponsor's brand of steel-belted radial tires. The sponsor was particularly interested in the utility functions of respondents who expressed interest in each of the test commercials.

The respondents considered tread mileage and price as quite important to their choice of tires. On the other hand, brand name did not play an important role (at least for the five brands included in the study). Not surprisingly, the most popular test commercial stressed tread mileage and good value for the money, characteristics of high appeal to this group. What was surprising was that this group represented 70% of the total sample.

This particular study involved the preparation of 25 profiles. Again, the researchers sorted cards into seven categories. The 25 profiles, also constructed according to an orthogonal array, represented only one twenty-fifth of the 625 possible combinations.

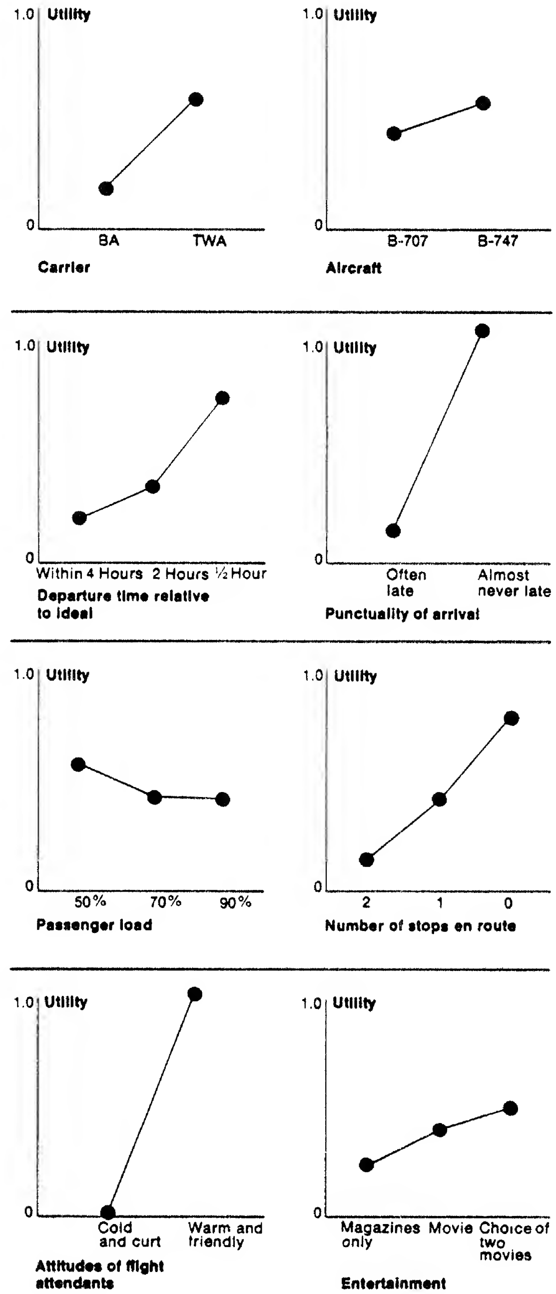


Figure 3–3. Utility Functions for Air Travelers to Paris

Potential Uses of Conjoint Measurement

The three preceding studies only scratched the surface of marketing problems in which conjoint measurement procedures can be used. For example, consumer evaluations can be obtained on:

New product formulations involving changes in the physical or chemical characteristics of the product

Package design, brand name, and promotional copy combinations

Pricing and brand alternatives

Verbalized descriptions of new products and services

Alternative service designs

Moreover, while the three preceding examples emphasized preference or likelihood-of-purchase orderings, any explicit judgmental criterion can be used. For example, alternatives might be ordered by any of these criteria:

Best value for the money

Convenience of use

Suitability for a specified type of consumer or for a specified end use

Ruggedness, distinctiveness, conservativeness, and other ''psychological images''

Designing Bar Soaps

In one recent study researchers related the psychological imagery of physical characteristics of actual bars of soap to end-use appropriateness; this study was conducted for the laboratory and marketing personnel of a diversified soap manufacturer.

While the designing of a bar of soap—by varying weight, size, shape, color, fragrance type and intensity, surface feel, and so on—may seem like a mundane exercise, the fact remains that a cleverly positioned bar soap (for example, Irish Spring) can rapidly become a multimillion-dollar enterprise. Still, the extent of knowledge about the importance of such imagery is woefully meager. The researchers formulated actual bars of soap in which color, type of fragrance, and intensity of fragrance were constructed according to a design in which all possible combinations of the experimental factors appeared. All the other characteristics of the soap were held constant.

Respondents examined the soaps and assigned each bar to the end use that they felt best matched its characteristics—moisturizing facial soap, deep-cleaning soap for oily skin, woman's deodorant soap, or man's deodorant soap. The data were then analyzed by conjoint measurement techniques, leading to a set of physical functions for each of the characteristics.

The study showed that type of fragrance was the most important physical variable contributing to end-use appropriateness. Rather surprisingly, the type of fragrance (medicinal) and color (blue) that appeared best suited for a man's deodorant soap were also found to be best for the deep-cleaning soap, even though deep-cleaning soap had been previously classed for marketing purposes as a facial soap. On the other hand, fragrance intensity played a relatively minor role as a consumer cue for distinguishing among different end uses.

In brief, this study illustrated the feasibility of translating changes in various physical variables into changes in psychological variables. Eventually, more detailed knowledge of these psychological transformations could enable a laboratory technician to synthesize color, fragrance, shape, and so forth to obtain soaps that conjure up almost any desired imagery. Moreover, in other product classes—beers, coffees, soft drinks—it appears possible to develop a psychophysics of taste in which such elusive verbal descriptions as "full-bodied" and "robust" are given operational meaning in terms of variations in physical or chemical characteristics.

Verbalized Descriptions of New Concepts

In many product classes, such as automobiles, houses, office machines, and computers, the possible design factors are myriad and expensive to vary physically for evaluation by the buying public. In cases such as these, the researcher usually resorts to verbalized descriptions of the principal factors of interest.

To illustrate, one study conducted among car owners by Rogers National Research, Inc., employed the format shown in figure 3–4. In this case the researchers were interested in the effects of gas mileage, price, country of manufacture, maximum speed, roominess, and length on consumer preferences for new automobiles. Consumers evaluated factor levels on a two-at-a-time basis, as illustrated in figure 3–4. Market Facts, Inc., employs a similar data collection procedure.[5]

In the Rogers study it was found that consumer evaluations of attributes were highly associated with the type of car currently owned and the type of car desired in the future. Not surprisingly, gas mileage and country of manufacture were highly important factors in respondent evaluations of car profiles. Somewhat surprising, however, was the fact that even large-car owners (and those contemplating the purchase of a large car) were more

concerned with gas economy than owners of that type of car had been historically. Thus, while they fully expected to get fewer miles per gallon than they would in compact cars, they felt quite strongly that the car should be economical compared to others in its size class.

Organizations as Consumers

Nor is conjoint measurement's potential limited to consumer applications. Evaluations of supply alternatives by an organizational buyer are similar to benefits sought by the consumer. Thus, one can argue, these evaluations are among the most important inputs to industrial marketing strategy.

As an illustration, the management of a clinical laboratory was concerned with the problem of how to increase its share of laboratory test business. It had a study conducted to assess how physicians subjectively value various characteristics of a clinical laboratory in deciding where to send their tests.

Each physician in the study received 16 profiles of hypothetical laboratory services, each showing a different set of characteristics, such as reliability of test results, pick-up and delivery procedures, convenience of location, price range of services, billing procedures, and turnaround time. Utility functions were developed for each of these factors. On the basis of these results the management of the laboratory decided to change its promotion by emphasizing a number of convenience factors in addition to its previous focus on test reliability.

Marketing Strategy Simulations

We have described a variety of applications of conjoint measurement, and still others, some in conjunction with the other techniques outlined in Appendix 3A, could be mentioned.[6] What has not yet been discussed, and is more important, is the role that utility measurement can play in the design of strategic marketing simulators. This type of application is one of the principal uses of conjoint measurement.

As a case in point, a large-scale study of consumer evaluations of airline services was conducted in which consumer utilities were developed for some 25 different service factors such as on-ground services, in-flight services, decor of cabins and seats, scheduling, routing, and price. Moreover, each utility function was developed on a route (city-pair) and purpose-of-trip basis.

As might be expected, the utility function for each of the various types of airline service differed according to the length and purpose of the flight. However, in addition to obtaining consumers' evaluations of service pro-

What is more important to you?

There are times when we have to give up one thing to get something else. And, since different people have different desires and priorities, the automotive industry wants to know what things are most important to you. We have a scale that will make it possible for you to tell us your preference in certain circumstances – for example, gas mileage vs speed. Please read the example below which explains how the scale works – and then tell us the order of your preference by writing in the numbers from 1 to 9 for each of the six questions that follow the example.

Example:
Warranty vs. price of the car

Procedure:
Simply write the number 1 in the combination that represents your first choice. In one of the remaining blank squares, write the number 2 for your second choice. Then write the number 3 for your third choice and so on from 1 to 9.

Years of warranty

Price of car	3	2	1
$3,000	1		
$3,200			
$3,400			

Step 1 (Explanation)

You would rather pay the least ($3,000) and get the most (3 years). Your first choice (1) is in the box as shown.

Years of warranty

Price of car	3	2	1
$3,000	1		
$3,200	2		
$3,400			

Step 2

Your second choice is that you would rather pay $3,200 and have a 3-year warranty than pay $3,000 and get a 2-year warranty.

Years of warranty

Price of car	3	2	1
$3,000	1	3	
$3 200	2		
$3,400			

Step 3

Your third choice is that you would rather pay $3,000 and have a 2-year warranty than pay $3,400 and get a 3-year warranty.

Years of warranty

Price of car	3	2	1
$3,000	1	3	6
$3,200	2	5	8
$3,400	4	7	9

Sample:

This shows a sample order of preference for all possible combinations. Of course, your preferences could be different.

For each of the six questions below, please write in the numbers from 1 to 9 to show your order of preference for your next new car.

Miles per gallon

Price of car	22	18	14
$3,000			
$3,200			
$3,400			

Miles per gallon

Maximum speed	22	18	14
80 mph			
70 mph			
60 mph			

Miles per gallon

Length	22	18	14
12 feet			
14 feet			
16 feet			

Miles per gallon

Roominess	22	18	14
6 passenger			
5 passenger			
4 passenger			

Miles per gallon

Made in	22	18	14
Germany			
U S			
Japan			

Price of car

Made in	$3,000	$3,200	$3,400
Germany			
U.S			
Japan			

Figure 3–4. A Two-at-a-Time Factor Evaluation Procedure

files, the researchers also obtained information concerning their *perceptions* of each airline (that is, for the ones they were familiar with) on each of the service factors for which the consumers were given a choice.

These two major pieces of information provided the principal basis for developing a simulation of airline services over all major traffic routes. The purpose of the simulation was to estimate the effect on market share that a change in the service configuration of the sponsor's services would have, route by route, if competitors did not follow suit. Later, the sponsor used the simulator to examine the effect of assumed retaliatory actions by its competitors. It was also able to use it to see what might happen to market share if the utility functions themselves were to change.

Each new service configuration was evaluated against the base-period

configuration. In addition, the simulator showed which competing airlines would lose business and which ones would gain business under various changes in perceived service levels. Thus, in addition to single, ad hoc studies, conjoint measurement can be used in the ongoing monitoring (via simulation) of consumer imagery and evaluations over time.

Prospects and Limitations

Like any new set of techniques, conjoint measurement's potential is difficult to evaluate at the present stage of development and application. Relatively few companies have experimented with the approach so far. Capability for doing the research is still concentrated in a relatively few consulting firms and companies.

Conjoint measurement faces the same kinds of limitations that confront any type of survey, or laboratory-like, technique. First, while some successes have been reported in using conjoint measurement to predict actual sales and market share, the number of applications is still too small to establish a convincing track record at the present time.

Second, some products or services may involve utility functions and decision rules that are not adequately captured by the models of conjoint measurement. While the current emphasis on additive models (absence of interactions) can be shifted to more complex, interactive models, the number of combinations required to estimate the interactions rapidly mounts. Still, little is known about how good an approximation the simpler models are to the more elaborate ones.

Third, the essence of some products and services may just not be well captured by a decomposition approach that assumes that the researcher can describe an alternative in terms of its component parts. Television personalities, hit records, movies, or even styling aspects of cars may not lend themselves to this type of reductionist approach.

While the limitations of conjoint measurement are not inconsequential, early experience suggests some interesting prospects for measuring consumer tradeoffs among various product or service characteristics. Perhaps what is most interesting about the technique is its flexibility in coping with a wide variety of management's understanding of consumers' problems that ultimately hinge on evaluations of complex alternatives that a choice among products presents them with.

Notes

1. R. Duncan Luce and John Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement." *Journal of Mathematical Psychology,* February 1964, p. 1.

2. The first marketing-oriented paper of conjoint measurement was by Paul E. Green and Vithala R. Rao, "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research,* August 1971, p. 355.

3. A nontechnical discussion of this special class of designs appears in Paul E. Green, "On the Design of Experiments involving Multiattribute Alternatives," *Journal of Consumer Research,* September 1974, p. 61.

4. As an illustration, see Joseph B. Kruskal, "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data," *Journal of the Royal Statistical Society,* Series B, March 1965, p. 251.

5. Richard M. Johnson, "Trade-Off Analysis of Consumer Values," *Journal of Marketing Research,* May 1974, p. 121.

6. Paul E. Green and Yoram Wind, *Multiattribute Decisions in Marketing: A Measurement Approach* (Hinsdale, Ill.:Dryden Press, 1973).

Appendix 3A: Other Techniques for Quantifying Consumers' Judgments

Conjoint measurement is the latest in an increasing family of techniques that psychometricians and others in the behavioral and statistical sciences have developed to measure persons' perceptions and preferences. Conjoint measurement can often be profitably used with one or more of the following:

Factor Analysis

Factor analysis in marketing research has been around since the 1940s. However, like all the techniques to be (briefly) described here, factor analysis did not reach any degree of sophistication or practicality until the advent of the computer made the extensive computations easy to carry out. A typical input to factor analysis consists of respondents' subjective ratings of brands or services on each of a set of attributes *provided by the researcher.* For example, a sample of computer systems personnel were asked to rate various computer manufacturers' equipment and services on each of the 15 attributes shown in figure 3A–1.

The objective of factor analysis is to examine the commonality across the various rating scales and find a geometric representation, or picture, of the objects (computers), as well as the attributes used in the rating task. As noted in figure 3A–1, International Business Machines (IBM) was ranked highest on virtually all attributes while Xerox (XDS), a comparatively new entrant at the time of the study, National Cash Register (NCR), and Central Data Corporation (CDC) were not perceived as highly as the others with regard to the various attributes of interest to computer users.

The tight grouping of the attribute vectors also suggests a strong "halo" effect in favor of IBM. Only in the case of price flexibility does IBM receive less than the highest rating, and even here it is rated a close second. Thus as figure 3A–1 shows, factor analysis enables the researcher to develop a picture of both the things being rated (the manufacturers) and the attributes along which the ratings take place.

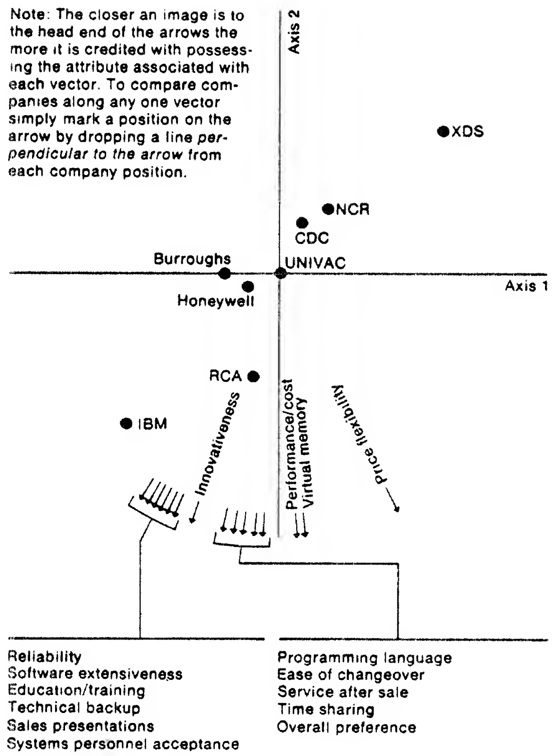


Figure 3A-1. Factor Analysis of Average Respondent Ratings of Eight Computer Manufacturers' Images on Each of Fifteen Attributes

Perceptual Mapping

A somewhat more recent technique—also abetted by the availability of the computer—is perceptual mapping. Perceptual mapping techniques take consumer judgments of *overall* similarity or preference and find literally a picture in which objects that are judged to be similar psychologically plot near each other in geometric space (see figure 3A-2). However, in perceptual mapping the respondent is free to choose *his own* frame of reference rather than to respond to explicitly stated attributes.

The perceptual map of the 11 automobiles shown was developed from consumers' judgments about the relative similarity of the 55 distinct pairs of cars that can be made up from the 11 cars listed. The dimension labels of *luxurious* and *sporty* do *not* come from the technique but rather from

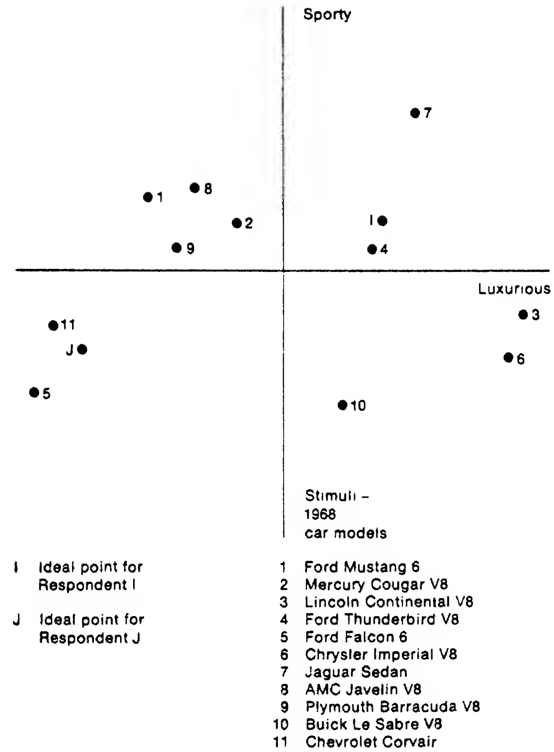


Figure 3A–2. Perceptual Mapping of Respondents' Judgments of the Relative Similarity of Eleven Cars and Two Respondents' Preference Orderings

further analysis of the map, once it is obtained from the computer. Ideal points I and J are shown for two illustrative respondents and are fitted into the perceptual map from the respondents' preference judgments. Car points near a respondents' ideal point are preferred to those farther away. Thus respondent I most likes Ford Thunderbird, while respondent J most likes Chevrolet Corvair. In practice, data for several hundred respondents might be used to find regions of high density for ideal points.

Cluster Analysis

Still another way to portray consumers' judgments is in terms of a hierarchical tree structure in which the more similar a set of objects is perceived to be, the more quickly the objects group together as one moves from left to

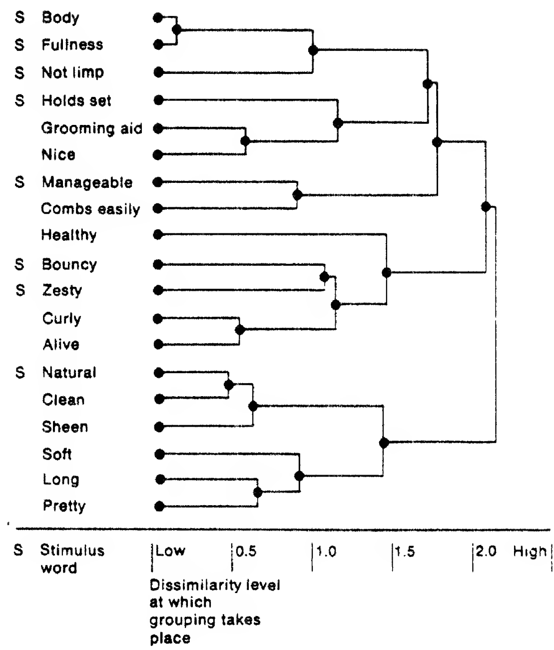


Figure 3A–3. Hierarchical Cluster Analysis of Nineteen Phrases Evoked in a Free-Association Task Involving Women's Hair Shampoos

right in the tree diagram. Thus the words *body* and *fullness* are perceived to be the two most closely associated of all of the descriptions appearing in figure 3A–3 that characterize hair. Note further that smaller clusters become embedded in larger ones until the last cluster on the right includes all 19 phrases. The words in this example were based on respondents' free associations to a set of 8 stimulus words. The researchers assumed that the more a stimulus evoked another word, the more similar they were.

Relationship to Conjoint Measurement

These three methods are best noted for their complementarities—both with each other and with conjoint measurement. Factor analysis and perceptual mapping can be used to measure consumers' perceptions of various products and services, while conjoint measurement can be used to quantify how consumers trade off some of one attribute to get more of another. Cluster analysis can be used in a variety of ways, either as a comparison technique

for portraying the similarities of various objects or as a basis for grouping people with common perceptions or preferences. In short, all these techniques can—and frequently are—applied in the *same* study. As such, their combined use can heighten different aspects of the same general types of input data.

4 New Computer Tools for Product Strategy

Paul E. Green and
J. Douglas Carroll

It is rather ironic that research and development's imagination and technological virtuosity in creating new products often outstrip its ability to predict their success in the marketplace. When one takes into account recent trends toward shorter product lifespans and increasing foreign competition, the future outlook for new-product successes appears even more risky.

It is no wonder that many companies are beginning to turn toward new-product design and testing procedures that can incorporate consumer preferences at early stages in the product's formation—well before the costly outlays of test marketing and regional introduction have to be made. Probably the most exciting of these new tools are the *product design optimizers* that have appeared on the marketing scene over the past year or two.

Product design optimizers are based on the simple idea that consumers respond to a *bundle* of product/image attributes, by trading off some things to get others. If the product strategist can control the attribute levels of the new product or product-concept description that the consumer evaluates, then the strategist can eventually tease out the relationship between overall preference and each of the attributes (including their interactions) that contribute to total evaluation.

This basic idea is not new. As early as 1951 statisticians were developing what has become known as *response-surface methodology* to explore the connections between controllable process variables and the yields of various chemical plants.[1] Once one finds this functional relationship, another technique—a process optimizer—is used to find the maximum yield, consistent with various constraints on the ranges in which the contributory variables (for example, temperature, chemical concentration) must lie. As far back as the early 1960s, these methods were being used in such diverse applications as improving the flakiness of pie crusts, reducing the costs of antibiotics, and improving the whiteness of bleached cotton.

However, what now appears new is the large-scale extension of these production-based ideas to products and services in which the *consumer* plays the role of "transducer"—in effect, a process that transforms the experimenter's product formulations or verbal concepts into overall evalua-

tion responses. In addition, other new techniques—specifically, multidimensional scaling[2] and conjoint analysis[3]—have also evolved to enhance the applicability of response-surface methodology to the twin problems of product positioning and market segmentation.

The purpose of this chapter is to describe what product design optimizers do and how they are being applied to various problems in the design of consumer and industrial products and services. We first discuss some of the limitations of a popular technique—conjoint analysis—that is already in use in a large number of companies and consulting firms. The basic ideas of product design optimizers are then introduced and described in the context of several large-scale applications. We conclude with a brief discussion of how this methodology may be extended in the future.

Conjoint Analysis and Its Limitations

Conjoint analysis has been on the marketing scene for about a decade. Hundreds of business applications have been made, and virtually all the larger marketing consulting firms have the capabilities for applying these methods. Conjoint analysis studies typically entail the following steps.

First, the researcher assembles a set of product attributes and levels. For example, if one were doing a study of resort-hotel preferences, one might select the attribute-level descriptions shown in table 4–1.

Table 4–1
Attribute-Level Descriptions of Resort Hotels

Food quality	*Nightlife/entertainment*
1. Superb	1. Much
2. Good	2. Some
3. Fair	3. Hardly any
Sightseeing places	*Chances for meeting new people*
1. Many	1. High
2. Some	2. Medium
3. Hardly any	3. Low
Outdoor sports	*Total trip cost*
1. Many	1. Inexpensive
2. Some	2. Average
3. Hardly any	3. Very expensive

FOOD
Superb

SIGHTSEEING PLACES
Hardly any

OUTDOOR SPORTS
Hardly any

NIGHTLIFE/ENTERTAINMENT
Some

CHANCES FOR MEETING NEW PEOPLE
Medium

TOTAL COST OF TRIP
Inexpensive

3

Overall Rating__________

Figure 4–1. Sample Stimulus Card Used in Preference Task

Next, the researcher selects from the $3^6 = 729$ possible combinations of the attribute levels of table 4–1 a small set (in this case only 18) that still enables one to estimate all average or main-effect utilities on an uncorrelated basis. A set of 18 prop cards, similar to the one shown in figure 4–1, is prepared, and the respondent is asked to sort the cards on a scale board containing usually seven to nine compartments, ranging from *like least* to *like most.*

Then, by using one of the conjoint-analysis algorithms, each person's utility function is found. These functions also enable the researcher to *estimate* responses to all the 711 remaining combinations, assuming that a simple additive part-worth model applies. Figure 4-2 shows one respondent's utility function. As noted, to find the utility for any specified combination, one sums its part-worths.

Finally, the individual utility functions are entered into a computerized choice simulator, along with some client-provided descriptions of various resort hotels in a particular geographic area (for example, Miami Beach). The computer simulator then plays back the share of choices that each hotel profile would receive if each respondent chose the hotel with the highest utility for him/her.

As simple as this method is, many firms have applied it successfully to problems in product and service positioning. However, there are some rather serious limitations to this procedure.

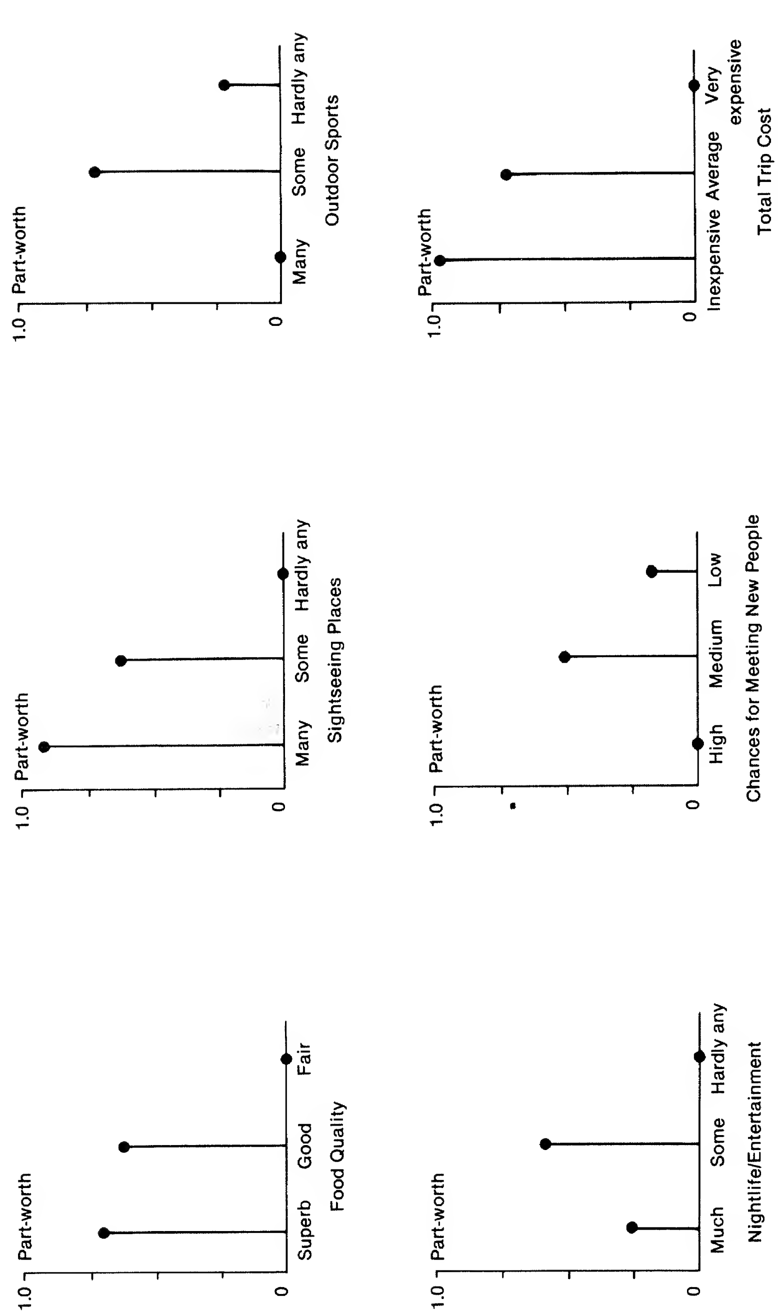


Figure 4–2. One Respondent's Utilities for Resort Hotel Attributes. (To find the utility for any combination, simply add its vertical-scale part-worths.)

First, interactions among pairs of attributes are not measured. For example, it may be the case that one's liking for a resort hotel offering superb-quality food *and* many sightseeing places is valued considerably more than the sum of their separate main-effect utilities. Interaction effects can be particularly strong in the design of food and beverage products or in other cases where sensory and aesthetic judgments are important.

Second, the researcher (and the client) have no idea whether the few test products that are tested in the simulator on an ad hoc basis are anywhere near the *best* product that could be designed, where "best" can be explicitly defined in terms of market share, profits, or any other measure of interest to the client.

Third, the researcher and client have no good data on how sensitive the response measure (for example, profits) would be to departures from the best levels of each attribute or what the best levels of other attributes would be if, for policy reasons, one or more attributes had to be fixed at some nonoptimal level.

Last, the traditional conjoint-analysis approach is static in that it assumes that the market instantaneously adjusts to the presence of a new item in some existing array of choices. No guidance is provided on how sales or profits may behave over the various stages in the new-product introduction period.

New-product design optimizers combine elements of conjoint analysis with some of the methodology of response-surface methodology to provide a considerably more rigorous and complete portrayal of consumer inputs to the new-product planning process.

Product Design Optimizers

The basic notion underlying response-surface methodology (one of the building blocks of product design optimizers) is depicted graphically in figure 4–3. Assume that only two attributes—the availability of outdoor sports and nightlife/entertainment activities—are involved. Moreover, each is assumed to be able to vary more or less continuously from none at all to some large number of activities. For the hypothetical respondent (no relation to the one in figure 4–2), whose function is depicted in figure 4–3, the utility for an increasing number of outdoor sports increases to reach a maximum of about five activities; then it starts to decline. Similarly, in the case of nightlife/entertainment activities, the utility reaches a maximum at about four activities and then declines very slowly. The principal idea of response-surface methodology is to model a customer's evaluative responses as a surface and then find the highest-utility point on the surface.

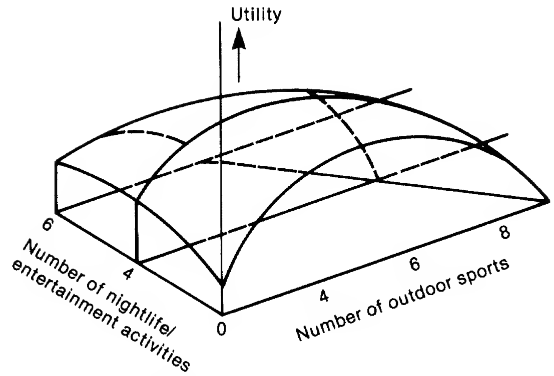


Figure 4-3. A Hypothetical Utility Response Surface

Real problems are complicated by many additional features. In general, there may be more than a dozen different attributes, not simply two. The surface may be quite irregular with twists and valleys in it. The utility for some attributes may rise continuously over the whole permissible range of the attribute rather than peaking and then dropping off. Some attributes, such as brand name or color, may be inherently categorical rather than continuous, and in that case there can be a different surface for each combination of the categorical variables.

In short, the simple geometric picture of figure 4-3 provides only a conceptual idea of what the product design strategist has to do, namely, perform a type of mathematical hill climbing to find the optimal value of the response function, within the technologically set ranges of the design variables.

A complete approach to product design optimization should include computer programs for accomplishing at least the following steps:

1. *Designing the experiment:* picking out of possibly thousands of possibilities those relatively few attribute-level combinations that are most efficient in estimating both main-effect utilities and selected pairwise-attribute interactions. In general, these designs lead to independent (uncorrelated) estimation of all contributory utilities.
2. *Data collection:* techniques that can obtain consumer responses to a limited number of attribute combinations but are still capable of estimating utility functions (at least partially) at the individual level.
3. *Modeling the choice simulator:* computer procedures for both simulating consumer choices and determining the mathematical relationship between the total sample's or specified segment's choices and the product-design variables that are under the control of the product strategist.

4. *Optimizing the model:* computer techniques for finding the optimum of the response function, given one's response to either categorical variables (as in figure 4–2), continuous variables (as in figure 4–3), or, more commonly, a mixture of the two types of variables.
5. *Sensitivity analysis:* exploring the effect on optimal response by departing from the best setting of each variable, or set of variables.
6. *Target market:* having found the best new-product positioning, one can then enter this formulation into the choice simulator and find *who* chooses it over existing products. Since respondents are also tagged in terms of their demographics or other background variables, one can obtain the best initial (target) segment for the new entry. This information provides help for designing advertising copy and selecting appropriate advertising vehicles.
7. *Sales/profit forecasting:* examining the time path of new-product trying and repeat purchase in terms of the source of the new business and customer brand-switching dynamics.

Despite their relative newness, product design optimizers have already permeated the sensory-testing field. CompuServe in Columbus, Ohio, and MPI Sensory Testing in New York City offer similar services (based on response-surface methodology) for finding optimal product designs. Applications have been made in such diverse product classes as asphalt, cheese, hand lotion, spaghetti sauce, rocket fuel, salad dressing, and synthetic rubber.[4] A number of marketing academics, both in the United States[5] and abroad,[6] have also developed approaches to product design optimization.

Possibly the most comprehensive of all these optimization approaches is the one called POSSE (*p*roduct *o*ptimization and *s*elected *s*egment *e*valuation), devised by the authors. POSSE combines conjoint-analysis modeling of consumer preferences with a variety of mathematical programming techniques for optimizing suitable objective functions (for example, contribution to overhead and profits), as modeled by a consumer-choice simulator. As of this writing, a set of twenty-one special-purpose programs has been developed by the first author (and his students) for implementing the various steps in the system. (See Appendix 4A for details).

How Product Design Optimization Works

In order to show as simply as possible how product design optimizers work, let us take a hypothetical example which is a composite, drawn from a set of real cases. Let us assume that a Japanese car manufacturer—call it Sioto—is interested in introducing a passenger van to the U.S. market. Preliminary research indicates that nine attributes (each at four levels) are important in most consumer's choices of a passenger van. Table 4–2 shows

Table 4–2
Product Attributes and Levels—Sioto Passenger Van

A. Cargo height, inches
1. 40
2. 46
3. 52
4. 58

B. Cargo length, inches
1. 86
2. 96
3. 106
4. 116

C. Cargo width, inches
1. 58
2. 62
3. 66
4. 70

D. Payload, pounds
1. 800
2. 1,200
3. 1,600
4. 2,000

E. Maintenance interval, miles
1. 3,000
2. 6,000
3. 9,000
4. 12,000

F. Standard features (no extra cost)
1. AM radio
2. FM stereo radio
3. FM stereo/CB radio
4. FM radio

G. Engine type, size, and gas mileage
1. 4-cylinder gasoline; 30 miles per gallon; $150 discount off base price
2. 6-cylinder gasoline; 18 miles per gallon; $0 discount off base price
3. 8-cylinder gasoline; 15 miles per gallon; $250 premium over base price
4. 6-cylinder diesel; 26 miles per gallon; $600 premium over base price

H. Type of interior
1. Standard; base price
2. Deluxe; $450 premium over base price
3. Plush; $700 premium over base price
4. Special; $1,000 premium over base price

I. Base price
1. $4,500
2. $4,800
3. $5,100
4. $5,400

the attribute levels needed to set up the experimental design. Some of these attributes, such as cargo height, represent discrete levels of an underlying continuous variable, while others, such as type of interior, are inherently categorical.

Experimental Design

The first step in applying product design optimization (for example, POSSE) is to develop the master experimental design that will be used to construct product-concept descriptions. In this case the computer developed a design that required the construction of only 64 attribute combinations

Table 4-3
A Portion of the Experimental Design for Selecting Stimulus Combinations

Stimulus	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
1	1	1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2	1	2
3	1	2	2	4	4	3	3	1	3
4	1	2	2	3	3	4	4	1	4
5	1	3	3	1	1	3	3	2	2
6	1	3	3	2	2	4	4	2	1
7	1	4	4	4	4	1	1	2	4
8	1	4	4	3	3	2	2	2	3
9	1	1	1	4	4	4	4	3	2
10	1	1	1	3	3	3	3	3	1
11	1	2	2	1	1	2	2	3	4
12	1	2	2	2	2	1	1	3	3
13	1	3	3	4	4	2	2	4	1
14	1	3	3	3	3	1	1	4	2
15	1	4	4	1	1	4	4	4	3
16	1	4	4	2	2	3	3	4	4
17	2	3	4	2	1	4	3	1	4
18	2	3	4	1	2	3	4	1	3
19	2	4	3	3	4	2	1	1	2
20	2	4	3	4	3	1	2	1	1
21	2	1	2	2	1	2	1	2	3
22	2	1	2	1	2	1	2	2	4
23	2	2	1	3	4	4	3	2	1
24	2	2	1	4	3	3	4	2	2

See table 4-2 for attribute and level descriptions.

(out of a possible total of $4^9 = 262{,}144$). Even at that, the design enables the researcher to estimate all the main effects and all two-way interactions among the first three attributes (height, length, and width), each uncorrelated with the rest.

Table 4-3 shows a portion of the design that was computer-generated. The computer also printed the specific eight (out of sixty-four combinations) that each respondent received, as well as the stimulus descriptions themselves. Each set of eight stimulus cards was also constructed so that every attribute level appeared an equal number of times *within* each respondent. Hence, even if some respondents tend to concentrate their responses at one end of the evaluative scale, this possible source of response bias can be isolated.

Data Collection

In this study a mail questionnaire was feasible, since the respondent only had to compare eight test profiles with a common control product, as illus-

ATTRIBUTE	TYPICAL CURRENT VAN	NEW VAN
Cargo Height	40 inches	40 inches
Cargo Length	96 inches	96 inches
Cargo Width	58 inches	62 inches
Payload	1,200 pounds	2,000 pounds
Maintenance Interval	6,000 miles	12,000 miles
Features	AM radio	FM Stereo/CM Radio
Engine Size/Type/ Gas Mileage *(cost versus base)*	6-cylinder gasoline/ 18 miles per gallon *(standard—no additional cost)*	8-cylinder gasoline 15 miles per gallon *($250 over standard)*
Interior *(cost versus base)*	Standard *(no additional cost)*	Standard *(no additional cost)*
Base Price	$4,500	$5,100 3

Figure 4–4. Comparison of Typical Current Brand and New Brand (Sample Card.)

trated in figure 4–4. Questionnaires were sent to respondents who were known (based on a telephone screening interview) to be in the market for a new van of this general type. For each comparison the respondent was asked to indicate his/her subjective likelihood of acquiring a new van with the characteristics described. In addition, data were collected on individual attribute-level desirability, attribute importances, and selected respondent demographics. The whole questionnaire was designed to take less than a half hour to complete. Four hundred fifty completed questionnaires were returned.

Modeling the Choice Simulator

The basic data, consisting of attribute-level desirabilities, attribute-importance weights, and full-profile evaluations, were then analyzed by other programs in the product-design methodology, resulting in an estimated individual utility function for each respondent. Next, a response "surface" was fitted to the data by introducing a set of 256 products (again drawn from a master experimental design), one at a time, into the choice simulator, along with a set of control (existing product) profiles.

In this case, the response variable was programmed to be expected contribution to the firm's profits and overhead if each of the test designs were introduced to the market. Another computer program was then used to reduce this mass of response data to a mathematical function that relates the total sample's design response to the controllable product-design variables of table 4–2.

Table 4–4
Optimal Values of Product Attributes as Found from Mixed Integer Non-Linear Programming Algorithms

A.	Cargo height 51 inches	*F.*	Standard features FM stereo/CB radio
B.	Cargo length 107 inches	*G.*	Engine type, size and mileage 6-cylinder diesel; 26 miles per gallon; $600 premium over base price
C.	Cargo width 68 inches	*H.*	Type of interior Deluxe
D.	Payload 1,100 pounds	*I.*	Price $4,850
E.	Maintenance level 6,245 miles		

Optimizing the Function

Other programs, utilizing mixed-integer nonlinear programming were then employed to find the combination of attribute levels that maximizes expected contribution. Table 4–4 shows the results. Since attributes *A, B, C, D, E,* and *I* (see table 4–2) are underlying continuous variables, the design optimizer is permitted to find a maximum that does not have to correspond with the discrete levels selected for the initial stimulus design evaluations.

Sensitivity Analysis

Once the optimal product design is found, other programs in the POSSE series can explore how sensitive the function is to departures from the optimum. For example, suppose technological constraints force Sioto to set the cargo width at 62 inches and the maintenance interval at 3,000 miles. What happens to the optimum under these conditions?

A subsequent run of the optimization programs showed the following new results:

Cargo height: 51 inches

Cargo length: 108 inches

Payload: 840 pounds

Standard features: FM stereo/ CB radio

Engine: 6 cylinder diesel: 26 miles per gallon; $600 premium

Interior: deluxe

Base price: $4,972

The technological constraints take their toll, however. The new optimum, under these *constrained* conditions, is 16.3 percent less than the unconstrained optimum whose profile appears in table 4–4.

Market Segmentation

Having found the optimal product for the total market, we can enter this profile into the choice simulator, and the researcher can now find the target-market segment that would be most attracted initially to the optimal product. The characteristics of this segment were:

Age: 21 to 32 years

Income: $17,600 to $24,300

Annual mileage: 9,400 to 13,650

Occupation: blue-collar

Current van owner? yes

Thus, at this point the strategist not only has identified the optimal product but also has found the best segment for initial promotional targeting.

Time-Path Forecasts

The methodology also contains a brand-switching algorithm that shows the time-path implications as customers switch to (and from) the new entry. Part of the program's output are gain-loss tables which show from what brands the new entry gains its share and to what brands it may eventually lose its share. In addition, the program projects market share (or sales, or profit) into the future so that the strategist can observe how these indicators build up as a consequence of the switching behavior.

The inputs to this part of the methodology consist of respondent's subjective likelihoods of switching to some other current brand or to the new entry. These likelihoods are separately estimated for each current brand.

While this particular computer program was not relevant for the van study, table 4–5 shows a small portion of the algorithm's output, as applied to a hypothetical market of three current brands and one new entrant. The initial shares are:

Brand 1:	65 percent	Brand 3:	10 percent
Brand 2:	25 percent	Brand 4 (new):	0 percent

Table 4–5
Sample Output of Brand-Switching Program

Brand-Switch Matrix Based on Average Incidence of Switch-out as Calculated from Old Brands

Last Bought/Next Bought	*1*	*2*	*3*	*4*[a]	*Total*
1	0.50	0.20	0.05	0.25	1.00
2	0.05	0.60	0.10	0.25	1.00
3	0.14	0.21	0.45	0.20	1.00
4	0.31	0.12	0.05	0.52	1.00

Projected Market Share of All Brands Including the New Brand (Based on Matrix of Average Switch-outs)

Time Period	*Brand 1*	*Brand 2*	*Brand 3*	*Brand 4*
1	0.65	0.25	0.10	0.00
2	0.35	0.30	0.10	0.25
3	0.28	0.30	0.10	0.32
4	0.26	0.29	0.11	0.34
5	0.26	0.29	0.11	0.34
6	0.26	0.29	0.11	0.34
7	0.26	0.29	0.11	0.34
8	0.26	0.29	0.11	0.34
9	0.26	0.29	0.11	0.34
10	0.26	0.29	0.11	0.34
Equilibrium share	0.26	0.29	0.11	0.34

[a]Brand 4 represents the new brand.

After ten purchase periods, however, we note that the new brand has captured 34 percent of the market, with the greatest loser being brand 1.

Some Actual Case Studies

The preceding hypothetical application of the product design optimizer is illustrative of the kinds of jobs these new computer tools can do. While the van study dealt with evaluations of product concepts, the optimizers can often be used profitably in the evaluation of actual products.

Designing an Optimal Frozen Pizza

In consumer evaluations of food, beverages, and other products where sensory impressions are important, verbal concepts do not adequately capture

the essence of the product. Such was the case in a study sponsored by a large national manufacturer of frozen pizzas.[7] While the firm's brand had enjoyed high popularity over the years, there were indications that new competitors would soon be chipping away at its share. Furthermore, from a market-segmentation standpoint, opportunities could exist for special pizza items aimed at selected target markets, such as teenagers or couples without children living at home.

The firm decided to apply product design optimization to this problem for two of its major products—the cheese pizza and the special. Illustratively, table 4-6 shows the ingredient levels that were selected for the special pizza. If all possible combinations were prepared, the researcher would have to test 157,464 combinations—clearly, an impractical task. In this case, however, a special experimental design was computer-generated that estimated all main effects and all two-attribute interactions for the first five attributes, with only 81 varieties of pizza.

Data were collected from several hundred respondents who tasted the pizzas in central kitchen testing facilities. Each respondent, over the course of a half-hour interview, tasted three of the eighty-one combinations—one at a time against a common control (the firm's current brand, unidentified, of course). In each case the respondent indicated which of the pair she/he most preferred and by how much—*slightly, somewhat,* or *very much.* Each test product was also rated on a set of sensory scales (for example, amount of sauce—*too much, just right, too little,* and so on). The respondent finally indicated the likelihood that the pizza would be tried, given a specified retail price. Background data on pizza consumption, demographics, and psychographics were also obtained.

It is interesting to note that in this type of application individual utility functions are bypassed. In this instance, the response measure is the likelihood that a randomly chosen respondent will select the test pizza over the (common) control, as a function of the test pizza's design variables. However, other procedures, known as *componential segmentation* methods,[8] can be used to estimate group differences in these probabilities, such as differences between teenagers and married adults.

The results of the study were useful in three ways. First, the experiment confirmed the sponsor's speculation that taste preferences differed markedly across market segments. Second, the experiment pointed out some attractive new-product design possibilities that would be expected to appeal to special segments, such as teenagers. Finally, the study showed just how important it was to have the capability to measure interactions between various pairs of attributes. In several instances the interactions were so strong (for example, amount of cheese versus type of crust) that failing to incorporate them in the product formulations would have led to extremely poor consumer acceptance.

Table 4-6
Ingredient Levels for Frozen-Pizza Experiment (Special Pizza Only)

Ingredient	*Number of Levels*
Type of crust	3
Type of cheese	3
Amount of cheese	3
Type of meat	3
Amount of meat	3
Type of sauce	3
Amount of sauce	3
Mushrooms?	2
Type of peppers	3
Olives?	2
Oil?	2
Price	3

Prototype: 81 varieties of test pizzas are formulated from the complete set of 157,464 possibilities.

Promoting a New Antihypertensive Drug

Sometimes new products are introduced in which product-design flexibility is limited, since the chemical characteristics are pretty much fixed. Most ethical drugs fall into this category. Still, the strategist may exercise control over such variables as pricing, dosage form, and advertising and promotion. Such was the situation in the case of a recently introduced drug for the treatment of hypertension. The sponsor of the study is a highly respected, international manufacturer of ethical drugs. The product under evaluation represented a new type of antihypertensive.

Seven attributes (see table 4-7), each at two to four levels, were pretested and found to be important determinants of choice insofar as this particular class of antihypertensives is concerned. A master design was set up that entailed thirty-two stimulus cards; this number was sufficient to estimate all main effects and selected attribute interactions. Each physician respondent received only four of the cards (with levels balanced within respondent) and was asked to indicate the likelihood on a 0 to 100 percent scale that he/she would prescribe the drug, given a specified patient description. (Four different patient histories, rotated so that each respondent received only one, were used.)

A mail sample stratified by physician specialty was used in this study; almost 500 returns were received. The physicians also supplied their preferences for current drugs, attribute importances and desirabilities, and

Table 4–7
Sample Stimulus Card for Antihypertensive-Drug Study

Attribute	*Current Antihypertensives*	*New Product*
Cardioselectivity	Noncardioselective (affects bronchi, stomach, and arteries, as well as the heart)	Cardioselective
Dosage regimen	Three times a day	Four times a day
Intrinsic sympathomimetic activity (ISA)	No significant ISA (does not increase heart rate nor drop blood pressure)	Same
Metabolism	Metabolized by body	Same
Absorption	Must be taken before meals	Not affected by meals
Renal blood flow	No change in renal blood flow	Decreases renal blood flow
Price	$0.50 daily dose	$1.00 per daily dose

various kinds of background data, including psychographics and attitudes toward antihypertensive drugs in general.

The product design optimizer produced a number of useful results, including the observation that physicians were price-sensitive to this drug and that utilities for the various attributes (see table 4–7) varied by physician segment. Four such segments were identified and profiled by physician specialty, attitudes, psychographics, and demographics.

What became immediately clear was that different promotional appeals would be needed for each segment; in turn, the segments could be accurately identified by specialty and a few other routinely available variables. Moreover, two important interactions dealing with preference and its relationship to absorption, renal blood flow, and metabolism were identified. These interactions were exploited in the preparation of promotional copy.

The sponsor, in collaboration with its advertising agency, acted on the study's findings by targeting its initial pricing and promotional copy for direct mail, journal advertisements, and detailing booklets along the lines suggested by the study. Variations on the main copy points were made to correspond with differences across the four segments' utilities.

The new drug was subsequently introduced and has received an excellent reception in the marketplace. While one cannot tell just how much of that success is due to the promotional-design recommendations, it is generally agreed that the segmented copy appeals have played a significant role.

One interesting postscript: As is so often the case in studies like this, the optimum combination of attribute levels for each segment was *not* one of

the thirty-two combinations used in the questionnaire evaluation. Rather, the optimum was inferred by maximizing the mathematical function linking choices with the (controllable) attributes. Since thousands of possible combinations are present in some studies (for example, the pizza project), it is comforting to know that one does not have to test all these—or even a large fraction of them—to find the optimum. This assumes, of course, that the model *is* a reasonably accurate description of reality.

The Planning of Rural Health-Care Facilities

Not all applications of new-product design strategies appear in the private sector of the economy. An excellent example of applying some of the principles described in this chapter is provided by a study whose objective was to expand rural primary health-care facilities so as to maximize the incremental social benefit to the community, subject to the usual constraints on cost. The study was carried out by Barnett R. Parker and V. Srinivasan.[9]

In this case the service attributes were type of facility (health center, doctor's office, nurse's office), travel time to facility, hours of operation, time to get an appointment, and average waiting time in the office. The respondents were drawn from the general environs of a community located in western New York State. Interviews were done on a personal, in-home basis; a sample size of 177 respondents was obtained.

Conjoint analysis was used to obtain the respondent's utilities. The utilities were then converted into a common metric that permitted the application of cost/benefit methods. While the particular problem did not lend itself to global optimization, the researchers were able to use a heuristic method that led to near-optimal community benefits within the cost constraints.

The Parker and Srinivasan study illustrates how many of the same techniques that are applicable to the private sector can also be appropriate for the evaluation of goods and services provided by not-for profit organizations.

Outlook for Product Design Optimizers

Unlike the hundreds of applications that have been made of the precursory technique, conjoint analysis, only a relatively few studies have been carried out using the more recent methodology of product design optimization. However, the future looks bright as more firms explore the potential of these new computer tools.

Product design optimizers are primarily concerned with the following strategic questions:

1. *Product Positioning*
   What is the most profitable new/modified product to make, consistent with the firm's current product line?
   What will be the switch-in and switch-out rates from/to competitive products?
   What is the optimal target market for this product?
2. *Market Segmentation*
   What market segments will optimize sales for the current product line?
   Given a specified market segment, what is the most profitable product for that segment?
3. *Competitive Reaction*
   Given the competitive introduction of a new product, what is the firm's best retaliatory strategy from a product standpoint?
   What is the firm's best retaliatory strategy from a market-segment standpoint?

While the current state of the art goes a long way toward helping to answer these questions, new developments are also expected to enhance the scope and power of the computer tools. Some examples are:

1. *Incorporation of multiple objectives*—in some problems the strategist may be dealing simultaneously with two or more response measures of interest, such as short-term profits, longer-term market share, and cannibalization of the firm's current products. What the strategist may wish to do is minimize departures from a set of desirable target levels. Research is proceeding on the adaptation of *goal-programming* methods for this case. (Goal programming is already in use in some of the sensory-testing approaches.)
2. *Product-line decisions*—in some applications the strategist may want to consider the concurrent introduction of two or more products, each designed to be attractive to different segments of the total market. There are some interesting subtleties associated with this problem, including the crosscompetition between the new products as well as their competitive draw from other firms' brands and the firm's older products.
3. *Retaliatory strategies*—situations frequently arise where a competitor introduces a new offering that upsets previous market relationships. How does one respond—by a retaliatory product, development of new segments, price reduction, or what? While research is proceeding on this class of problems, the complexities are considerable (but so are the stakes).
4. *Incorporating other strategic variables*—the ethical-drug example shows how product design optimization can be adapted to promotional

design. Still, additional research is needed on the question of how much to spend on advertising, distribution, and other policy variables. Current design optimizers emphasize product attributes, pricing, promotional content, and target segments. These models can be augmented by additional inputs that provide at least rough guidance on how much to spend for promotion and distribution and how to allocate these funds among sales regions, media, and the like.

5. *Public-service activities*—as the Parker and Srinivasan study illustrates, the potential for product/service design optimization is not limited to profit-oriented organizations. The methodology has potential for measuring differences in utilities for various constituencies—government, corporations, trade unions, consumers—and then designing public programs that attempt to balance these often conflicting objectives. Research is now underway in developing methods for assessing the relative weights of the constituencies when conflicts need to be resolved.

While the outlook for continued development and application of product design optimizers is good, one item should not be neglected, namely, the *validation* of predictions made by these techniques. Validation is a difficult and costly undertaking—one that management is not always interested in doing, particularly since the validation itself may turn out to be ambiguous. Nonetheless, some opportunities exist for at least reasonably controlled conditions, such as in-store experiments. And, however difficult the task, this research is vital to the longer-term value of the current methodology and the enhancements that are planned for the future.

Notes

1. G.E.P Box and K.B. Wilson, "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society* **13B**(1951): 1–45.
2. J. Douglas Carroll, "Individual Differences and Multidimensional Scaling," in *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences,* vol. 1, edited by R.N. Shepard, K. Romney, and S.B. Nerlove (New York: Seminar Press, 1972), pp. 105–155; Allan D. Shocker and V. Srinivasan, "A Consumer-Based Methodology for the Identification of New Product Ideas," *Management Science* (February 1974): 921–937.
3. For a recent survey, see Paul E. Green and V. Srinivasan, "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research* (September 1978) pp. 103–23.
4. *Marketing Communications,* November 1978, pp. 53–56.

5. Fred S. Zufryden, "ZIPMAP—A Zero Integer Programming Model for Market Segmentation and Product Positioning," *Journal of the Operational Research Society,* 30(1979): 63-70.

6. Sönke Albers and Klaus Brockhoff, "A Procedure for Product Positioning in an Attribute Space," *European Journal of Operational Research* 1(1977): 230-238.

7. This study, and the one immediately following, were carried out by the first author, using the POSSE methodology.

8. For an explanation of componential segmentation, see Paul E. Green and Wayne S. DeSarbo, "Componential Segmentation in the Analysis of Consumer Trade-offs," *Journal of Marketing,* (Fall 1979): 83-91.

9. Barnett R. Parker and V. Srinivasan, "A Consumer Preference Approach to the Planning of Rural Primary Health-Care Facilities," *Operations Research* (September-October 1976): 991-1025.

References

Albers, S. 1979. "An Extended Algorithm for Optimal Product Positioning." *European Journal of Operational Research* 3, 222-231.

Albers, S. and K. Brockhoff. 1977. "A Procedure for New Product Positioning in an Attribute Space." *European Journal of Operations Research* 1, 230-238.

———. 1979. "A Comparison of Two Approaches to the Optimal Positioning of a New Product in an Attribute Space." *Zeitschrift fur Operations Research,* Band 23, 127-142.

Bachem, A. and H. Simon. 1977. "Optimal Product Positioning in an Attribute Space." Working paper, University of Bonn, Bonn, Germany.

Benson, Purnell H. 1966. "Analysis of Irregular Two Dimensional Distributions of Consumer Buying Choice." *Journal of Marketing Research* 3 (August): 278-288.

———. 1965. "Fitting and Analyzing Distribution Curves of Consumer Choices." *Journal of Advertising Research* 5 (March): 28-34.

Carroll, J.D. 1972. "Individual Differences and Multidimensional Scaling." In *Multi-dimensional Scaling.* Theory, vol. 1. Edited by R.N. Shepard, A.K. Romney, and S.B. Nerlove. New York: Seminar Press, 105-155.

Coombs, C.H. 1964. *A Theory of Data.* New York: John Wiley.

Gavish, B.; D. Horsky; and K. Srikanth. 1979. "Optimal Positioning of a New Product." Working paper MERC 79-05, Graduate School of Management, University of Rochester.

Green, Paul E. 1975. "Marketing Applications of MDS: Assessment and Outlook." *Journal of Marketing* 39 (January): 24–31.

Green, Paul E. and Frank Carmone. 1970. *"Multidimensional Scaling and Related Techniques in Marketing Analysis.* Boston: Allyn and Bacon.

Green, Paul E. and V. Srinivasan. 1978. "Conjoint Analysis in Consumer Research: Issues and Outlook. *The Journal of Consumer Research* 5 (September): 103–123.

Green, Paul E. and Y. Wind. 1973. *Multiattribute Decisions in Marketing.* Hinsdale, Ill.: The Dryden Press.

Hauser, J.R. and P. Simone. 1979. "Product Realization: Selection of Physical Features and Price to Achieve a Profit Maximization Perceptual Position." Working paper, Northwestern University, August.

Johnson, R.M. 1971. "Market Segmentation: A Strategic Management Tool." *Journal of Marketing Research* 8 (February): 13–18.

Kuehn, A.A. and R.L. Day. 1962. "Strategy of Product Quality." *Harvard Business Review* 40 (November–December): 100–110.

McFadden, D. 1976. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement* 5 (Fall): 363–390.

Morgan, N. and J. Purnell. 1969. "Isolating Openings for New Products in a Multidimensional Space." *Journal of the Market Research Society* 11 (July): 245–266.

Parker, B.R. and V. Srinivasan. 1976. "A Consumer Preference Approach to the Planning of Rural Primary Health Care Facilities." *Operations Research* 24(September): 991–1025.

Pekelman, D. and S. Sen. 1974. "Mathematical Programming for the Determination of Attribute Weights." *Management Science* 20 (April): 1217–1229.

Pessemier, E.A. 1975. "Market Structure Analysis of New Product and Market Opportunities." *Journal of Contemporary Business* (Spring): 35–67.

Pessemier, E.A. and P. Root. 1973. "The Dimensions of New Product Planning." *Journal of Marketing* 37 (January): 10–18.

Shocker, A.D.; J. May; and D. Sudharshan. 1978. "Design of New Product Concepts: A Comparison of Optimizing Approaches." Paper presented at Joint National ORSA/TIMS Conference, Los Angeles, (November).

Shocker, A.D. and V. Srinivasan. 1974. "A Consumer-Based Methodology for the Identification of New Product Ideas." *Management Science* 20 (February): 921–937.

———. 1979. "Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review." *Journal of Marketing Research* 16 (May): 159–180.

Shugan, S.M. and V. Balachandran. 1977. " Mathematical Programming

Model for Optimal Product Line Structuring." Working paper, University of Rochester.

Srinivasan, V. 1976. "Network Models for Estimating Brand Images." Research Paper No. 344. Graduate School of Business, Stanford University, December.

Srinivasan, V. and A.D. Shocker. 1973a. "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments." *Psychometrica* 38 (December): 473–493.

———. 1973b. "Linear Programming Techniques for Multidimensional Analysis of Preference." *Psychometrica* 38 (September): 337–369.

Stefflre, V. 1969. "Market Structure Studies: New Products for Old Markets and New Markets (Foreign) for Old Products." In *Application of the Sciences to Marketing.* Edited by Bass, King, Pessemier. New York: John Wiley 251–268.

———. 1971. *New Products and New Enterprises: A Report on an Experiment in Applied Social Science.* Irvine, Calif.: University of California, March.

Urban, G.L. 1975. "PERCEPTOR: A Model for Product Positioning." *Management Science* 21 (April): 858–871.

Zeleny, M. 1976. "The Attribute-Dynamic Attitude Model (ADAM)." *Managment Science* 23 (September): 12–26.

Zufryden, F.S. 1977. "A Conjoint Measurement-Based Approach for Optimal New Product Design and Market Segmentation." Working paper presented at the AMA Marketing Science Workshop at the University of Pittsburgh in November 1977, Graduate School of Business, University of Southern California.

———. 1979. "ZIPMAP—A Zero-One Integer Programing Model for Market Segmentation and Product Positioning." *Journal of Operational Research Society* 30, 63–70.

Appendix 4A: The POSSE Programs

As mentioned in the body of this chapter, POSSE currently has twenty-one special-purpose programs that are used to implement the system. One thing that we have become convinced of, since initiating POSSE and applying it to real-world problems, is that no single model or algorithm is sufficient to deal with the complexity and variety of problems that the applications-oriented researcher faces. For this reason, POSSE contains several different programs with broadly similar objectives. Moreover, work is still going on in extending the series to cope with new situations that arise as the methodology is applied to more diverse product/service classes. What seems to be emerging is a set of program modules that can be flexibly applied to a variety of problem situations.

Listing of POSSE Programs

The POSSE suite of programs contains a mixture of readily available, general-purpose programs (for example, regression, ANOVA, gradient-free polynomial optimization) and special programs that have been written *de novo*. Only this later set of programs is described here.

The (current) programs in the POSSE series are listed under seven main headings:

1. Experimental design—DESIGN, LABEL, ORTHOTEST
2. Choice simulators—CHOSIM, BTLSIM, DUESIM, DUEALL
3. Categorical optimization—QUALO, QUALIN
4. Polynomial optimization—QUADMO, MOGUL, INDEF, RANDO
5. Sensitivity analysis—RIDGE, TOLCO 1, TOLKO K, TOLCO M
6. Regression—MONREG, THRESH
7. Forecasting—SWITCH, BENEMAX

Each program is described (in varying degrees of detail) accroding to the preceding format.

Experimental Design Programs

The POSSE methodology makes use of various kinds of fractional factorial designs, including resolution 5 designs (where all main effects and two-fac-

tor interactions are orthogonally estimable), compromise designs (where all main effects and selected two-factor interactions are orthogonally estimable), and orthogonal main-effects plans. In most POSSE applications, respondents receive only four to nine product profiles; the profiles are usually selected so that attribute levels are balanced within respondent.

The three programs that are employed in this phase of POSSE are DESIGN, LABEL, and ORTHOTEST.

The DESIGN Program

The DESIGN program is the cornerstone of POSSE's stimulus design process. This is the program that prepares the actual master experimental layout from which the stimulus props are eventually prepared. The DESIGN program can be used to prepare designs drawn from the following classes:

- Orthogonal main-effects designs
- Resolution 4 designs, where all main effects can be orthogonally estimated, free of all two-factor interactions
- Compromise designs, where all main effects and selected two-factor interactions can be orthogonally estimated
- Resolution 5 designs, where all main effects and all two-factor interactions can be orthogonally estimated

The basic factorials from which the preceding classes of designs can be constructed are the following:

2^5:	32 rows and 31 columns
2^6:	64 rows and 63 columns
2^7:	128 rows and 127 columns
3^3:	27 rows and 13 columns
3^4:	81 rows and 40 columns
3^5:	243 rows and 121 columns
4^3:	64 rows and 21 columns
4^4:	256 rows and 85 columns
5^3:	125 rows and 31 columns

In each case, the full design (including all interaction terms) is constructed and can be printed out, at the user's option.

Alternatively, the user may ask for only selected columns of the overall design. As an additional feature, the user may ask that some of the entries in one or more of the selected columns be collapsed, for example, that all three-levels be collapsed to two-levels or that all three-levels and two-levels be collapsed to zero-levels. (This collapsing feature extends the versatility of the program to handle asymmetric designs).

In all cases DESIGN prints out the complete legend for the full table and the requested columns, appropriately numbered. Also, a record of the collapsing instructions is printed.

The LABEL Program

Once the master design has been developed (through the use of the DESIGN program), the user can employ a second program, LABEL, to prepare the actual stimulus props. LABEL accepts as input the master design, the attribute names, and the attribute-level names.

The program then prints out the prop descriptions in sequence, including the particular stimuli that each respondent should receive, assuming that a respondent-blocking option is used.

The ORTHOTEST Program

ORTHOTEST is the last program in the series. ORTHOTEST takes an experimental design and tests its columns for orthogonality, based on the following proportionality test. Let

n = number of stimuli (rows in plan)

r = number of attributes (columns in plan)

x_{ip}, x_{jp} = the ith (jth) level of the pth attribute ($p = 1, 2, \ldots, r$)

$n_{i.}$ = number of times level i of attribute p occurs

$n_{.j}$ = number of times level j of attribute $p'(\neq p)$ occurs

n_{ij} = number of times level i of attribute p occurs with level j of attribute p'

Hence

$$\sum_j n_{ij} = n_{i\cdot} \qquad \sum_i n_{ij} = n_{\cdot j} \qquad \sum_{i,j} n_{ij} = n$$

In order to satisfy orthogonality, the following condition is both necessary and sufficient:

$$n_{ij} = n_{i.}n_{.j}/n$$

for all $p \times p'$ pairs of columns and all i,j pairs of levels.

Choice Simulators

POSSE employs different kinds of simulation programs that are used to compute the share of choices under different test-product descriptions. The simulators—CHOSIM, BTLSIM, DUESIM, and DUEALL—utilize different approaches to the modeling of product choice. However, all the programs exhibit the common objective of computing share of choices received by a specific test product in the context of one or more control products.

All four programs carry the same design specifications of handling up to forty attributes, with up to ten levels of each attribute, and simulations of up to ten products. All the simulation programs are currently designed for up to 1,000 respondents, but can be easily redimensionalized upward. In short, all four simulation programs are designed to handle large-scale problems if necessary.

The CHOSIM Program

The CHOSIM simulator has the capability of performing four different types of simulation: deterministic perceptions with individual utility functions, stochastic perceptions with individual utility functions, deterministic perceptions with group utilities and an associated error term, and stochastic perceptions with group utilities and an associated error term. In the latter two cases, the group utilities can represent a particular segment of interest to the researcher or the total sample, as the case may be.

Deterministic Perceptions with Individual Utility Functions

The deterministic/individual utility-function case is the one that most current simulators consider. In this case the product configurations under test are assumed to be perceived without error. All utility functions are assumed to be estimated without error.

Three types of choice models are implemented in CHOSIM:

1. Choice based on maximum utility; this is the most commonly used option.
2. The Bradley-Terry-Luce mode; this is defined as

$$P_j = \frac{u_j}{\sum_{j=1}^{J} u_j}$$

where P_j is the probability that the jth option is chosen and u_j is the utility for the jth option.
3. The logit choice model; this is defined as

$$P_j = \frac{e^{u_j}}{\sum_{j=1}^{J} e^{u_j}}$$

where P_j and u_j are defined as before and e is the base of the natural-log system.

In addition, the program prints out the average rank number, as associated with the maximum utility option.

Stochastic Perceptions with Individual Utility Functions

This option is similar to the one above with the exception that the user reads in a set of probability functions, one for each attribute, that reflect respondents' (assumed) perceptions that the product will have each of K levels. For example, although a new automobile might be designed to deliver 24 miles per gallon, its perceived (or, in some cases, actual) mileage may be 18, 22, 24, 26 miles per gallon, and so on.

Perception dispersions will depend on the inherent variability in the attribute being modeled, the lack of explicitly available information on the actual attribute level, and persons' misperceptions about the attribute, whether correct information is available or not.

Since this option utilizes probability distributions for attribute levels, its results will differ from the preceding case (unless, of course, all the probability mass is centered on a single level, per attribute). The choice models used in this option are the same as those described above.

Deterministic Perceptions with Group Utilities with Associated Error

In this option the user assumes that all the product configurations being simulated are perceived without error. However, rather than having N individual utility functions, this option assumes a single group-level utility function with an error term that reflects individual variation around the group utilities. This error term is assumed to be normally distributed with zero mean and standard deviation computed from the mean square error associated with the regression model, from which the group utility function is obtained.

The choice models used in this option are the same as those described above. However, only one utility function—the group average—is read in. (Hence this case is reasonable only when the probabilistic utility option is used.)

Stochastic Perceptions with Group Utilities with Associated Error

In this option two sources of error are introduced—perceptual error and utility error, as reflected by departures of individual utilities from the group average. As such, different runs of the same data set can be expected to produce different results, reflecting the Monte Carlo nature of the design. Accordingly, it is generally useful to make several runs of the same data set in order to see how varied the results associated with the "best" option are.

The choice models used in this option are the same as those described above.

The BTLSIM Program

The BTLSIM model shares some of the features of CHOSIM, namely the ability to consider stochastic perceptions as well as deterministic ones and the ability to utilize group as well as individual utility functions.

However, the choice model employed in BTLSIM differs in two major respects from CHOSIM. First, BTLSIM assumes that the response variable of interest is a subjective probability that the jth new option will be chosen, as compared to maintaining the status quo. (This is a common type of response variable and one that figures prominently in the POSSE methodology.) As such, BTLSIM assumes that the utilities that serve as inputs to the program were obtained initially from a group-level regression in which the response variable is

$$\psi_j = \ln\left[\frac{P_j}{1 - P_j}\right]$$

where ψ_j is the logit, P_j is the initially stated subjective probability, and ln is the base of the natural-log system.

BTLSIM finds the predicted probability $\hat{P}_j$ by the inverse transformation

$$\hat{P}_j = \frac{1}{1 + e^{-\psi_j}}$$

In this way all predicted probabilities will fall in the 0–1 interval.

The Bradley-Terry-Luce choice model is then employed to estimate first-choice probabilities in situations involving three or more alternatives. First, the parameter Π_j is computed from

$$\Pi_j = \frac{\hat{P}_j}{1 - \hat{P}_j}$$

for each option. Then, the following substitution is made:

$$\hat{P}_j^* = \frac{\Pi_j}{\sum_{j=1}^{J} \Pi_j + 1}$$

and the status quo option P_0^* is found by subtraction:

$$\hat{P}_0^* = 1 - \sum_{j=1}^{J} \hat{P}_j^*$$

At this point, the group-level probabilities have been obtained.

A full account of this adaptation of the BTL model is given by Green, Carroll, and DeSarbo.[1]

A Monte Carlo procedure is then used to simulate individual choice, assuming that each individual in the group respects the group-level probabilities. However, an error term is included in the group-level utilities, similar to the structure used in CHOSIM.

Alternatively, a set of individual elevation constants $b_0^{(i)}$ can be read in for each respondent $i = 1, 2, \ldots, I$ which leads to a separate probability function for each respondent; this is also sampled via the Monte Carlo procedure.

As in CHOSIM, the BTLSIM program also handles stochastic or deterministic perceptions, at the user's option. The choice model that is finally implemented consists of selecting the option associated with the highest choice probability on that particular trial.

The DUESIM Program

DUESIM is a deterministic version of CHOSIM that utilizes a hybrid utility model. By "hybrid" utility model is meant that the utilities of some attribute levels are estimated by one method (for example, self-explication) while the utilities of other attribute levels are estimated by some other method (for example, conjoint analysis).

DUESIM also permits the researcher to introduce brand-specific utilities. Moreover, all analyses can be carried out at the segment level as well as for the total (weighted) sample.

A full account of the model underlying DUESIM is given by Green, Goldberg, and Montemayor.[2]

The DUEALL Program

DUEALL is constructed along lines similar to DUESIM, namely, a hybrid utility model that uses a deterministic (maximum-utility) choice rule. However, in DUEALL the objective is to introduce, in addition, a *set* of new product profiles, based on a fractional factorial design, that are used to find estimated share of first choices which are later fitted by a logit model, prior to optimization via QUALO or a combination of QUALO and one of the modified gradient or grid-search procedures (in the case of quantitative attributes) in the POSSE series.

It should be mentioned that DUEALL is easily modifiable to handle other objective functions that can be obtained from the simulator:

1. Average probability of first choice (as obtained by a BTL or logit-based rule)
2. Expected contribution to overhead and profit
3. Share of choices received by *any* of the firm's products (including existing brands)
4. Share of choices received from selected segments
5. Expected revenues received by the firm's leading competitors (a criterion that the user may wish to minimize)

and so on. Indeed, multiple measures can be programmed, and techniques utilizing nonlinear goal programming can be used.

Choice of Simulators

CHOSIM, BTLSIM, and DUESIM are employed on conjoint-analysis problems where the researcher plans to terminate the analysis at the choice-simulation stage. In contrast, the purpose of DUEALL is to provide a dependent variable (for example, share of first choices as a function of test-product profile) for subsequent parameterization and optimization by other programs in the POSSE series.

Categorical Optimization

The POSSE system is flexible enough to handle the most general case of mixed attributes—categorical and continuous, including two-factor interactions. In the case of categorical attributes, the two-factor interactions are based on dummy-variable coding, while in the case of continuous variables, cross products are used. (If categorical attributes interact with continuous attributes, conditional optima are first found for the continuous variables, given specified levels of the discrete variables; one then optimizes over the discrete variables.)

Optimization is carried out with separate sets of programs, by taking advantage of function separability. In the case of categorical variables, QUALO and QUALIN are employed.

The QUALO Program

The QUALO program is designed to find optimal test products in the case where all attributes are categorical to begin with or where we are interested in only a relatively small number of discrete levels of underlying continuous attributes. In either case, the test product is considered to be the cartesian product of discrete sets of attribute levels.

The criterion variable in QUALO can be any of the measures (for example, share of first choices) mentioned earlier. The criterion is obtained by regressing the appropriate output of DUEALL on the dummy design variables that describe the simulated test product. Depending on the problem, logit transformations of the dependent variable may be used, prior to running the regression. The results of the regression are a set of part-worths that express the contribution of each attribute level's main effect (and pairs of attributes' two-way interactions) to the criterion variable.

Of course, if any attribute does not interact with any of the rest, its maximum part-worth is found simply by inspection. More realistically, if the attributes can be partitioned so that all two-way interactions are *within-set* interactions, then the optimization can proceed separately (and indepen-

dently) for each subset. In this manner QUALO is capable of handling large problems of thirty attributes or more.

Any specific run of QUALO is currently programmed for up to twenty attributes, each with two to nine levels; all 190 two-way interaction tables (or any proper subset) can be handled. The optimum is found via implicit enumeration.

In addition to the optimum test product, QUALO can print out the second highest, third highest, and so on, up to the number specified by the user (a maximum of fifteen).

The QUALIN Program

The QUALIN program is a version of QUALO that deals with individual (rather than segment-based) utility functions. Similar inputs are used, but in this case each function pertains to a specific individual. In QUALIN the objective is to find the test profile that maximizes the number of consumers for whom the new product is the first choice (that is, exhibits the highest utility). An explicit enumeration technique is used to find the optimum.

Constrained Optima

In both QUALO and QUALIN the user is allowed to prespecify one or more attribute levels and then find the best test product, conditional on these prespecified levels.

Polynomial Optimization

In many applications of POSSE, the attributes will be composed of underlying continuous variables. If such is the case, the user might wish to optimize some polynomial function of these variables, under linear constraints. In the most usual case, the function chosen will be *quadratic,* and the constraints will simply represent allowable ranges on each variable.

A number of programs—QUADMO, MOGUL, INDEF, and RANDO—have been developed to handle the case of quadratic functions. In general, all these programs yield only a local optimum; hence it is a prudent procedure to employ several different programs and several different starting configurations within each program, in order to assure oneself that a satisfactory optimum has been obtained. All four of the programs in this set can handle up to forty independent variables, with up to eighty upper and lower (range) constraints.

The QUADMO Program

QUADMO is an optimization program that exhibits the primary advantages of relatively fast convergence and low susceptibility to local optima in the case of functions with "saddle-point" solutions. QUADMO is restricted to general quadratic functions of the form

$$y = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{kk}x_k^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots$$
$$+ a_{k-1,k}x_{k-1}x_k + b_1x_1 + b_2x_2 + \cdots + b_kx_k + c$$

Alternatively, in matrix notation, we can write

$$y = \mathbf{x}'\mathbf{A}\mathbf{x} + \mathbf{b}'\mathbf{x} + c$$

where $\mathbf{x}$ is a column vector, $\mathbf{A}$ is a matrix of squared terms and cross products ($a_{ij}/2$ for $i \neq j$), $\mathbf{b}$ is a column vector of coefficients for the linear terms, and c is a constant.

For example, given the quadratic function

$$y = -x_1^2 + 3x_2^2 + x_1x_2 + x_1 + 2$$

we can write

$$y = (x_1 \quad x_2)\overset{\mathbf{A}}{\begin{bmatrix} -1.0 & 0.5 \\ 0.5 & 3.0 \end{bmatrix}}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \overset{\mathbf{b}}{(1.0 \quad 0)}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 2$$

Next, we define $\nabla\mathbf{x}$, the *gradient* of $y = f(\mathbf{x})$, as a vector of first-order partial derivatives:

$$\nabla\mathbf{x} = 2\mathbf{A}\mathbf{x} + \mathbf{b}$$

Similarly, the *Hessian* of $y = f(\mathbf{x})$ is defined as a matrix of second-order partial derivatives:

$$\mathbf{H} = 2\mathbf{A}$$

QUADMO utilizes a modified Hessian $\mathbf{H}^*$ that is obtained by singular-value decomposition of $\mathbf{H}$, followed by modification of the eigenvalues of $\mathbf{H}$, so that $\mathbf{H}^*$ is negative definite.

First, find the singular-value decomposition of $\mathbf{H}$:

$$\mathbf{H} = \mathbf{UBU}'$$

where $\mathbf{B}$ is diagonal and $\mathbf{U}'\mathbf{U} = \mathbf{I}$, the identity matrix. Next, compute $\mathbf{B}^* = \text{diag}\,(b_i^*)$ via the substitutions

$$b_i^* = \begin{cases} b_i & \text{if } b_i < 0 \\ -b_i & \text{if } b_i > 0 \\ -k & \text{if } b_i = 0 \end{cases}$$

where

$$k = \min_{b_i \neq 0} |b_i|$$

The modified Hessian is then defined as

$$\mathbf{H}^* = \mathbf{UB}^*\mathbf{U}'$$

Finding the Modified Gradient

The modified gradient in QUADMO is defined as

$$\begin{aligned} \mathbf{g}^*(\mathbf{x}) &\equiv (\mathbf{H}^*)^{-1}\, \nabla \mathbf{x} \\ &= (\mathbf{H}^*)^{-1}(2\mathbf{Ax} + \mathbf{b}) \\ &= 2(\mathbf{H}^*)^{-1}\mathbf{Ax} + (\mathbf{H}^*)^{-1}\mathbf{b} \\ &= \mathbf{Cx} + \mathbf{d} \end{aligned}$$

where

$$\mathbf{C} = 2(\mathbf{H}^*)^{-1}\mathbf{A} \qquad \mathbf{d} = (\mathbf{H}^*)^{-1}\mathbf{b}$$

Note that $\mathbf{C}$ and $\mathbf{d}$ are computed once and for all in a given run of QUADMO.

Defining the Change Vector

The change vector $\Delta\mathbf{X}^I$, associated with the Ith iteration, is defined as

$$\Delta \mathbf{X}^I = \begin{bmatrix} \Delta x_1^I \\ \Delta x_2^I \\ \vdots \\ \Delta x_k^I \end{bmatrix}$$

where (1) if $\ell_k < x_k^I < u_k$ (where ℓ_k and u_k are the respective lower and upper boundaries for x_k^I), then

$$\Delta x_k^I = -g_k^{\bullet}(\mathbf{x})$$

(2) if $x_k^I = \ell_k$, then

$$\Delta x_k^I = \max\ [0,\ -g_k^{\bullet}(\mathbf{x})]$$

and (3) if$x_k^I = u_k$, then

$$\Delta x_k^I = \min\ [0,\ -g_k^{\bullet}(\mathbf{x})]$$

The change in sign of $g^*(\mathbf{x})$ is done because we wish to maximize $y = f(\mathbf{x})$. Since $\mathbf{H}^{\bullet}$ has been earlier defined to be negative definite, we change the sign of $g^*(\mathbf{x})$ so that the step size will always be positive. Thus, $y = f(\mathbf{x})$ will never be decreasing over successive iterations of QUADMO.

Note that if the kth component of $\Delta\mathbf{x}^I$ is on one of the boundaries and $-\mathbf{g}^*(\mathbf{x})$ is attempting to take it outside the boundary, QUADMO will zero out this component so that $\mathbf{x}$ stays within the preset linear constraints (that is, the prespecified boundaries of each variable x_k).

Starting and Stopping Conditions

The algorithm starts out with a user-specified starting vector whose components are chosen to be within the user-specified boundaries of each component of $\mathbf{x}$. Furthermore, on each iteration the program prints the value of the function on the Ith iteration and the length of $\Delta\mathbf{x}^I$, defined as

$$\epsilon^* = \|\Delta\mathbf{x}^I\|$$

If $\|\Delta\mathbf{x}^I\| \leq \epsilon$, a user-specified stopping criterion, then the algorithm terminates. QUADMO also limits the number of total iterations to a preset maximum of fifty; however, this can be easily modified by the user.

The primary advantages of QUADMO over straight gradient procedures are its faster convergence and general ability to deal with functions whose quadratic forms are not necessarily negative definite.

The MOGUL Program

MOGUL also optimizes a quadratic function under constraints on the ranges of the independent variables. However, in MOGUL a three-phased algorithm is used. The first phase operates on that part of the Hessian whose eigenvalues are zero or positive. The second phase operates on that part whose eigenvalues are negative. The third phase is a straight-gradient procedure. In this way, one partitions the optimization space so that one always moves in the direction of maximization. In most other respects, MOGUL is very similar to QUADMO.

The INDEF Program

INDEF is a program that has been adapted from an original algorithm, proposed by James Bunch, of the University of California at San Diego, and Linda Kaufman, of Bell Laboratories, Murray Hill, New Jersey.[3] The Bunch and Kaufman algorithm was originally programmed by Linda Kaufman; INDEF is an adaptation of that program that has been designed to be compatible with QUADMO and MOGUL. Like QUADMO and MOGUL, INDEF is designed to handle the case of saddle-point and ridge-type surfaces. Moreover, the user can specify more general linear constraints on the independent variables than simply the range constraints that are incorporated in QUADMO and MOGUL.

The RANDO Program

RANDO is an optimization program that simply evaluates the quadratic function for a large set of random or selected (by experimental design) points. The random points are generated internally, while the user reads in design values, should this option be employed; grid search represents a special case of this model.

Not only is the maximum value printed (function value and optimal levels of the predictors), but also the user can specify the highest K solutions (where $K \leq 25$). Range constraints can be placed on all independent variables.

The main purpose of RANDO is to serve as a check on programs like

QUADMO, MOGUL, and INDEF that provide a systematic search for the optimum via gradient-type procedures.

Sensitivity Analysis

After an optimum is found by some program(s) in the POSSE series, the researcher may wish to explore the nature of the response "surface" around the optimum. Four programs—RIDGE, TOLCO 1, TOLCO K, and TOLCO M—are available for this purpose.

The RIDGE Program

RIDGE uses the technique of ridge analysis to explore a quadratic surface for *global* optima that fall on the circumference of hyperspheres of increasing radius from the center of the experimental design. Ridge analysis, as described in Hoerl,[4] is a useful diagnostic tool for determining how sensitive the response is to each of the independent variables as one moves away from the center of the experimental design via a series of concentric hemispheres of increasing radius. A more complete discussion of RIDGE, as used in the POSSE series, can be found in Green, Carroll, and Kedia.[5]

The TOLCO 1 Program

Tolco 1 is also designed for general quadratic functions. The TOLCO 1 program is applied after the optimal solution is computed by one of the quadratic programming algorithms. TOLCO 1 performs a type of sensitivity analysis for each variable, on a one-at-a-time basis. The user reads in the following input: the tolerance criterion ϵ, step size λ, vector of independent variables that maximizes the criterion variable y, general quadratic function (initially optimized), and upper and lower constant boundaries for each independent variable. The tolerance criterion ϵ can be expressed either in absolute terms or as a decimal fraction of y^{max}.

Suppose the user specifies ϵ to be 10 percent of y^{max} which, in turn is 13.333; then ϵ is 1.3333. Assume that the user also specifies each boundary as

$\ell_k \equiv$ lower boundary for kth independent variable; ℓ_k is set equal to 0

$u_k \equiv$ upper boundary for kth independent variable; u_k is set equal to 1.5

and the step size as $\lambda = 0.03$. Given these user-specified parameters, TOLCO 1 first considers x_1, the first independent variable. Its optimal value $x_1^{\max}$ will fall into one of three classes:

$$x_1^{\max} = \ell_k = 0$$
$$x_1^{\max} = u_k = 1.5$$
$$0 < x_1^{\max} < 1.5$$

If $x_1^{\max} = 0$, then TOLCO 1 proceeds to hold all other independent variables $x_2, x_3, \ldots, x_n$ at their optimal levels. It then increments on x_1 in steps of 0.03 and computes the value of y for each such step until y falls below $y^{\max} - \epsilon$, at which point the iterations on this variable stop.

If $x_1^{\max} = 1.5$, then TOLCO 1 holds all other independent variables at their optimal levels. It then decrements on x_1 in steps of -0.03 and computes the value of y for each such step until y falls below $y^{\max} - \epsilon$, at which point the iterations on this variable stop.

If $0 < x_1^{\max} < 1.5$, then two solutions are found. The program first moves x_1 to the left of $x_1^{\max}$ by decrementing in steps of -0.03 until either $y^{\max} - \epsilon$ or zero, the lower constraint, is reached. The program then moves x_1 to the right of $x_1^{\max}$ by incrementing in steps of 0.03 until either $y_1^{\max} - \epsilon$ or 1.5, the upper constraint, is reached. As such, two solutions are found for the case in which $x_1^{\max}$ is an interior point.

TOLCO 1 then proceeds to $x_2^{\max}$ and repeats the process, conditioned on all other variables $x_1, x_3, x_4, \ldots, x_n$ being held at their optimal values.

The result of all this is a step-by-step summary of how sensitive the objective function is to each of the independent variables and how far each of the variables can move from its optimal value before y becomes less than $y^{\max} - \epsilon$, the stopping value.

TOLCO 1 is dimensionalized to handle up to forty independent variables and eighty linear constraints (lower- and upper-boundary points); these are the same limits that are built into QUADMO, MOGUL, INDEF, and RANDO.

The TOLCO K Program

The TOLCO K program is a generalization of the TOLCO 1 program that enables the researcher to develop simultaneous tolerance intervals for $k (1 \leq k \leq n)$ independent variables. TOLCO K considers user-supplied change vectors (away from the optimal $\mathbf{x}^{\max}$ vector) and their impact on the criterion variable.

In order to get some idea of the mathematics underlying TOLCO K, consider the following problem. We assume that a researcher has solved for

the optimal vector $\mathbf{x}^{max}$ and wishes to see how sensitive the objective function is to departures from the optimum, assuming that these departures involve a *simultaneous* displacement of all k independent variables from their respective optima.

To consider this problem algebraically, we make the following definitions:

$y \equiv \mathbf{x}'\mathbf{A}\mathbf{x} + \mathbf{b}'\mathbf{x} + t$ denotes the objective function with boundaries ℓ_k and u_k for the kth variable; $\mathbf{A}$ denotes the matrix of squared terms and cross product, $\mathbf{b}'$ denotes the row vector of linear coefficients, and t denotes the constant term

$x_{\max} \equiv$ a vector $\mathbf{x}$ that maximizes $y = f(\mathbf{x})$

$y^{\max} \equiv$ maximum value of y

$\epsilon \equiv$ a user-specified value that indicates an allowable departure from $y^{\max}$

$\mathbf{d} \equiv$ a direction vector away from the optimum vector $\mathbf{x}^{\max}$

Next, we set up the following relation:

$$f(\alpha) = y(\mathbf{x}^{\max} + \alpha\mathbf{d})$$

We then expand in terms of α to get

$$\begin{aligned} f(\alpha) &= \mathbf{x}'^{\max}\mathbf{A}\mathbf{x}^{\max} + 2\alpha(\mathbf{d}'\mathbf{A}\mathbf{x}^{\max}) + \alpha^2\mathbf{d}'\mathbf{A}\mathbf{d}' + \mathbf{b}'\mathbf{x}^{\max} + \alpha\mathbf{b}'\mathbf{d} + t \\ &= y^{\max} + \alpha^2(\mathbf{d}'\mathbf{A}\mathbf{d}) + 2\alpha(\mathbf{d}'\mathbf{A}\mathbf{x}^{\max}) + \alpha(\mathbf{b}'\mathbf{d}) \end{aligned}$$

Hence

$$f(\alpha) = \alpha^2(\mathbf{d}'\mathbf{A}\mathbf{d}) + \alpha(2\mathbf{A}\mathbf{x}^{\max} + \mathbf{b})'\mathbf{d} + y^{\max}$$

or

$$f(\alpha) = K_1\alpha^2 + K_2\alpha + \mathbf{K}_3$$

where

$$\begin{aligned} K_1 &= \mathbf{d}'\mathbf{A}\mathbf{D} \\ K_2 &= 2(\mathbf{A}\mathbf{x}^{\max} + \mathbf{b})'\mathbf{d} \\ &= 2\mathbf{x}'^{\max}\mathbf{A}\mathbf{d} + \mathbf{b}'\mathbf{d} \\ K_3 &= y^{\max} \end{aligned}$$

We then set $f(\alpha)$ equal to the following:

$$f(\alpha) = K_1\alpha^2 + K_2\alpha + K_3 = y^{\max} - \epsilon$$

or

$$f(\alpha) = K_1\alpha^2 + K_2\alpha + \epsilon = 0$$

We note that $f(\alpha)$ is simply a quadratic equation. When solved, this will generally result in two solutions:

$$\mathbf{x}_1 = \mathbf{x}^{\max} + \alpha_1\mathbf{d} \qquad \mathbf{x}_2 = \mathbf{x}^{\max} + \alpha_2\mathbf{d}$$

To examine the nature of the solution, we make the usual substitutions,

$$K_1 \equiv a \qquad K_2 \equiv b \qquad K_3 \equiv c$$

and compute the discriminant of the quadratic function $b^2 - 4ac$ with the possible outcomes

$$b^2 - 4ac \quad \begin{cases} > 0 \\ = 0 \\ < 0 \end{cases}$$

If $b^2 - 4ac = 0$ or $b^2 - 4ac < 0$, then the tolerance region is the whole feasible region of $\mathbf{x}$, and TOLCO K stops at this point.

However, assume that $b^2 - 4ac > 0$. If so, two cases are possible:

Case 1: All the components $\mathbf{x}_\ell$ ($\ell = 1, 2$) are in the feasible region. If so, $\mathbf{x}_\ell$ defines a tolerance boundary for $\mathbf{x}$.

Case 2: Not all the components of $\mathbf{x}_\ell$ ($\ell = 1, 2$) are in the feasible region. In this case we must find a nonnegative constant k that defines an $\mathbf{x}_\ell^*$ that *is* in the region. In general, $\mathbf{x}_\ell^*$ will be on a boundary or will be $\mathbf{x}^{\max}$ itself.

Thus, if case 2 occurs, we proceed as follows:

$$f(\alpha) = \mathbf{x}^{\max} + k\,\alpha_\ell\mathbf{d}_\ell \qquad k \geq 0$$

Next we make the following definitions:

$$z_j \equiv j\text{th component of } \mathbf{x}^{\max}$$

$$w_j \;=\; \alpha_\ell \text{ times the } j\text{th component of } \mathbf{d}, \text{ or } w_j = \alpha_\ell \mathbf{d}_{\ell j}$$

$$v_j \;\equiv\; \begin{cases} u_j & \text{if } w_j > 0 \\ \ell_j & \text{if } w_j < 0 \\ z_j & \text{if } w_j = 0 \end{cases}$$

Then we find

$$k = \min_j \frac{v_j - z_j}{w_j} \qquad \text{where } k \geq 0$$

and

$$\frac{v_j - z_j}{w_j} = \infty \qquad \text{if } w_j = 0$$

That is, we ignore all cases where $w_j = 0$.

In case 2 we then solve for

$$\mathbf{x}^{\bullet}_\ell = \mathbf{x}^{\max} + k\alpha_\ell \mathbf{d}_\ell$$

and it may turn out that $k = 0$, so that one of the solutions is $\mathbf{x}^{\max}$ itself.

In practice, several different values of **d** may be chosen (either systematically or randomly), and simultaneous tolerance contours are computed. Moreover, the researcher can examine the effect on $y = f(\mathbf{x})$ of varying only a subset of the independent variables while holding the remainder at their optimal levels. In short, TOLCO K is a highly flexible program for examining how far the independent variables can *simultaneously* vary from their optimal values while maintaining $y = f(\mathbf{x})$ within a tolerance value of ϵ from $y^{\max}$.

The TOLCO M Program

TOLCO M is an extension of TOLCO 1 that allows the user to examine the sensitivity of the optimum to changes in both categorical and continuous variables when each variable systematically departs from the optimal setting. Continuous variables are handled in the same way as described under TOLCO 1. In the case of categorical variables, each alternative level of a given variable is examined in terms of its effect on the optimum, with all other variables held at their optimal settings.

Regression

In POSSE various types of regression programs are employed in the fitting of either utility functions or other kinds of response functions. In addition to a variety of standard regression routines, the programs developed especially for POSSE are MONREG and THRESH.

The MONREG Program

MONREG is a monotonic regression program that is primarily used to fit utility functions in cases where the user assumes that only an ordinally scaled dependent variable is available. (The independent variables are assumed to be either intervally scaled or dummy variables.)

MONREG is based on a conventional regression program with a subroutine (due to Carroll and Chang) that computes a best-fitting monotonic function of the predicted values iteratively. As such, no stronger assumption than ordinal scaling is assumed for the dependent variable.

The THRESH Program

THRESH is an alternating least-squares program used to fit parameters in cases where the respondent is assumed to follow a threshold-type utility model. That is, for certain attribute levels, the respondent is assumed to behave as though the utility of these levels were so low that they effectively swamped all other utilities, leading to low total evaluations of the test profiles that exhibit these critically poor attribute levels.

An alternating least-squares procedure is used to iteratively estimate the large negative utilities so as to maximize the fit between the dependent variable and a linear combination of the utilities of all attribute levels contained in a specific profile. A multistage hybrid model is used to collect the data.

Forecasting Procedures

Several miscellaneous programs have been prepared to carry out various ancillary analyses. Only two of these—SWITCH and BENEMAX—are described here; these bear most directly on the POSSE system.

The SWITCH Program

The SWITCH program is a forecasting program in POSSE that helps answer such questions as: What is the future trend of market share, by

brand, if no new product is introduced? What is the future trend of customer gains and losses attributed to consumer switching? What is each brand's share at equilibrium? First, assume that a new product is introduced. POSSE can be used to compute an estimated probability of trying the new product. Moreover, data are also obtained on a posteriori basis (after exposure to the new-product description) on the likelihood of purchasing each of K current brands. From these data SWITCH can answer the following research questions:

1. What is the trend of try rate for the new products?
2. What is the trend of market share for the new product, given an average incidence of switching *out* of the new product to a competitive (current) product; a best case (lowest incidence) of switching out of the new product; a worst case (highest incidence) of switching out of the new product; and a user-supplied switch-out vector?
3. What is the gain-loss pattern associated with the new product's entry?
4. What are the equilibrium brand shares for the new product and the current brands?

The Model

The SWITCH model is based on some relatively straightforward notions from Markovian stochastic modeling. For what follows we assume a first-order Markov process in which the probability of switching from state i to state j depends only on state i and not on states prior to i.

We make the following definitions:

$$\mathbf{p}^{(1)} \equiv (p_1^{(1)}, p_2^{(1)}, \ldots, p_K^{(1)})$$

denotes the initial probability of a randomly chosen consumer in each of K states, where K is the number of current brands under consideration. The superscript refers to time period $t = 1$. Thus, the vector $\mathbf{p}^{(1)}$ can be thought of as an initial customer-share vector at time $t = 1$. (If desired, its components can reflect the relative buying quantities of each consumer.) And

$$\mathbf{A} \equiv \{a_{ij}\}_{K \times K}$$

denotes a $K \times K$ matrix of brand-switching probabilities, as estimated from the a priori constant-sum point assignment in POSSE (before introduction of the test item) over the K current brands. Since the questionnaire also records which brand the respondent purchased last, the initial state i of each respondent is known as well.

Market-Share Trend – Current Brands

Given $p^{(1)}$ and **A**, it is a simple matter to find customer share at time period $t = 1, 2, \ldots, T$ from the recursive relationship

$$p^{t+1} = p^t\mathbf{A} \qquad \text{for } t = 1, 2, \ldots, T$$

SWITCH prints out these shares of market, by time period, for $T \leq 40$, as specified by the user.

Gain-Loss Summary – Current Brands

To find the gain-loss customer summary, we make the following stipulations:

1. Assume that each row of **A** sums to 100.
2. Add the number of persons switching out of brand i to all other brands; record the total number of switch-outs.
3. Add the number of persons switching into brand i from all other brands; record the total number of switch-ins.
4. Subtract the number of switch-outs from the number of switch-ins to get the net gain or loss.
5. Repeat for each time period and record, by brand i, the number of switch-outs, the number of switch-ins, and the net gain (or loss).

Market-Share Trend—New-Brand Case

Inasmuch as the typical application of POSSE does not estimate the probability of switching *out* of the new brand to a current brand, some assumptions have to be made about this vector of probabilities in the switching matrix $\mathbf{A}^*$. (However, if POSSE is used in actual in-home-use product tests, one can obtain empirical estimates of this vector of probabilities.)

The first set of computations utilized by SWITCH assumes that the switch-out probability is the *average* experienced by current brands:

$$S_o = \sum_{i=1}^{K} (1 - a_{ii})/K$$

Hence, $1 - S_o$ denotes brand loyalty to the new item.

The specific switch-out probabilities are found by allocating S_o across the K brands in proportion to their market shares:

$$\mathbf{q}^{(1)} = (q_1^{(1)}, q_2^{(1)}, \ldots, q_K^{(1)})$$

at time period $t = 1$. This set of assumptions provides a $(K + 1)$st row of the matrix $\mathbf{A}^*$, and $q_{K+1}^{(1)}$ is assumed to be zero.

In the *best-case* version, the switch-out probability for the new brand is found from

$$S_b = 1 - \max_{i \epsilon K}[a^*_{ii}]$$

over $i = 1, 2, \ldots, K$. That is, we find the current brand that exhibits the least switch-out and assume that the test brand can equal this performance. Allocations over the K current brands are handled in the same way as noted above.

In the *worst case,* the switch-out probability for the new brand is found from

$$S_w = 1 - \min_{i \epsilon K}[a^*_{ii}]$$

over $i = 1, 2, \ldots, K$. We then proceed in the same way as noted above.

SWITCH also allows the user to read in a switch-out vector of one's own choosing.

The BENEMAX Program

BENEMAX is a special-purpose program that is used in cases where interest centers on the optimal set of product benefits to promote. BENEMAX utilizes simple input data regarding respondents' ideal sets of product benefits, conditioned on the number of benefits that may be advertised.

BENEMAX is both a forecasting and an optimization procedure. For each set of possible product benefits that could be advertised, BENEMAX finds the weighted number of respondents for whom various numbers of desired benefits are reached by the advertisement. An optimization feature enables the user to find a mix of benefits that maximizes one of several optimal criterion functions, for example, the number of respondents for whom K or more benefits from the consumer's ideal set are present.

BENEMAX is currently programmed for up to 1,000 respondents and up to fifteen product benefits. However, it can easily be redimensionalized to handle more benefits or larger sample sizes. Also BENEMAX has the

capability of incorporating goal-programming features when multiple-criterion functions are appropriate.

Notes

1. Paul E. Green, J. Douglas Carroll, and Wayne S. DeSarbo, "Estimating Choice Probabilities in Multiattribute Decision Making" (Working Paper, University of Pennsylvania, May 1979).
2. Paul E. Green, Stephen M. Goldberg, and Mila Montemayor, "A Hybrid Utility Estimation Model for Conjoint Analysis" (Working Paper, University of Pennsylvania, January 1980).
3. James R. Bunch and Linda Kaufman, "Indefinite Quadratic Programming" (Computing Science Technical Report No. 61, Bell Laboratories, Murray Hill, N.J., June 1977).
4. Arthur E. Hoerl, "Ridge Analysis," *Chemical Engineering News* 60(50):67–71.
5. Paul E. Green, J. Douglas Carroll, and Pradeep K. Kedia, "Ridge Analysis in Optimal Product Design" (Working Paper, University of Pennsylvania, August 1979).

5 A Comparison of Two Logit Models in the Analysis of Qualitative Marketing Data

David Flath and *E.W. Leonard*

Increasing attention is being directed to statistical models for explaining categorical choices by survey respondents, such as their adoption or nonadoption of a new service. These choices are represented mathematically by qualitative variables which assume one of only two possible values. "Logit"-type models have been shown to be appropriate for analysis of such data and their mathematical properties have been well studied, first by Berkson (1955) and more recently by Nerlove and Press (1973). But practitioners of marketing research are not yet fully informed of these results, despite a fine start by Green, Carmone, and Wachspress (1977). The following discussion is intended to further that effort.

A maximum likelihood logit model is described and shown to be superior to a logit-type model presented by Green, Carmone, and Wachspress (1977), the "weighted least squares logit."[1] Also, interpretation of results of logit-type models is discussed, with special attention to interaction effects.

An Example

The estimation problems addressed here are understood best in the context of a specific example. The results of a randomized survey conducted by Leonard (1978) are shown in table 5-1. Retail bank customers were surveyed as to the characteristics thought related to probability of their having adopted a recently introduced automatic teller service. From the responses, qualitative variables were constructed:

$$\text{ADOPT} = \begin{cases} 1 & \text{if adopted automatic teller service} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{AGE} = \begin{cases} 1 & \text{if older than 35 years} \\ 0 & \text{if otherwise} \end{cases}$$

Reprinted from D. Flath and E.W. Leonard, "A Comparison of Two Logit Models in the Analysis of Qualitative Marketing Data," *Journal of Marketing Research* 16 (November 1979):533–538, published by the American Marketing Association.

Table 5-1
Proportion of Respondents Adopting Automatic Teller Service, Related to Age, Income, and Mobility
(*sample size = 100*)

	Age ≤ 35 Years		*Age > 35 Years*	
Income per Year	*Nonmobile*	*Mobile*	*Nonmobile*	*Mobile*
≤$15,000	0.222 (9)	0.692 (13)	0.000 (11)	1.000 (1)
>$15,000	0.500 (12)	0.809 (21)	0.500 (28)	0.200 (5)

Note: Number of observations in each cell is shown in parentheses. The original data set included many more cells than the eight shown here. For a discussion and analysis of the complete set, see E.W. Leonard, "Life Style Measures and Segmentation in Retail Banking" (Ph.D. dissertation, Oklahoma State University, 1978).

$$\text{INCOME} = \begin{cases} 1 & \text{if annual income greater than \$15,000} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{MOBILITY} = \begin{cases} 1 & \text{if lived in at least 3 different cities in last 10 years} \\ 0 & \text{otherwise} \end{cases}$$

"ADOPT" is the criterion variable in a dependence structure with "AGE," "INCOME," and "MOBILITY" as predictors.

The traditional marketing approach to analyzing data sets like these is first to group observations into cells. Then a statistical model is defined which is based on the means of observations in each cell. An alternative approach is to base the statistical model on the individual observations themselves, rather than on the cells of observations. In the example used here, the traditional approach leads to a weighted least squares regression procedure, whereas the latter implies maximum likelihood estimation. In either case the criterion variable is "probability of adoption."

Logit-Type Models

If the criterion variable is constrained to lie between zero and one, use of a linear or log-linear function is precluded. In this case, the response function is a probability function by definition. Two functions which have been used in models with qualitative dependent variables are the normal and the logistic. The first defines the "probit" class of models and the other the

"logit" models. Because the logit models present much simpler computation problems, attention in this article is restricted to that class.[2]

The logistic *cdf* is defined as

$$y = f(\mathbf{x}) = \frac{1}{1 + e^{-\boldsymbol{\beta}\mathbf{x}}} \tag{5.1}$$

where

$$\boldsymbol{\beta}\mathbf{x} = \beta_0 + \beta_1 X_1 + \beta_k X_k$$

The function represents an S-shaped surface with inflection always occurring at $Y = 1/2$. Changes in β_0 shift the surface laterally whereas changes in $\beta_1, \ldots, \beta_k$ affect its dispersion (see figures 5-1 and 5-2). The logistic function can represent a fairly close approximation to the normal, which is (heuristically) an advantage. If y falls in the closed interval [0, 1], equation 5.1 can be rewritten

$$\log \frac{y}{1 - y} = \boldsymbol{\beta}\mathbf{x} \tag{5.1a}$$

The left side of equation 5.1*a* is called the "logit" of y. All logit-type models share this general form and its considerable advantages over ANOVA, cross-tabulation, or other conventional techniques. In logit

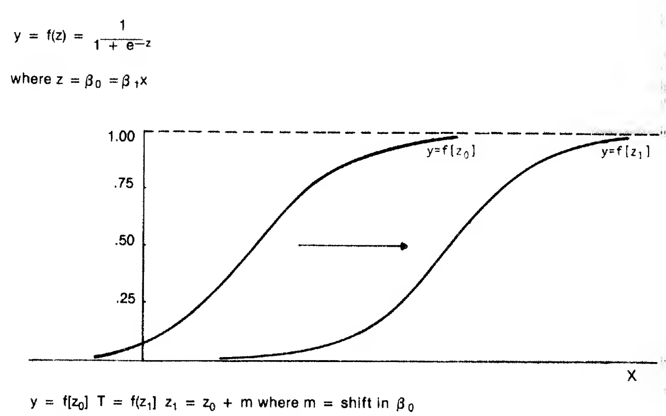


Figure 5–1. Effect of Change in the Constant Parameter on the Mapping of Logistic Function

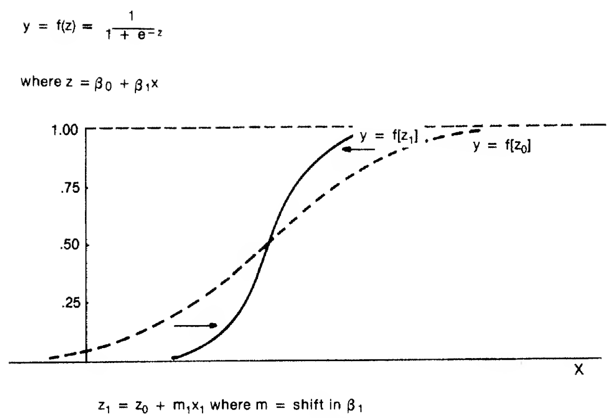


Figure 5-2. Effect of Change in a Slope Parameter on the Mapping of Logistic Function

models, the predicted values of the criterion variable have intuitive appeal as probability estimates, and estimates of the coefficient vector ($\boldsymbol{\beta}$) can be used in calculating direction and magnitude of marginal effects.

Weighted Least Squares Logit

A weighted least squares logit (WLS logit) model has been defined by Berkson (1955) and applied to marketing data by Green, Carmone, and Wachspress (1977). Let "P_j" represent the proportion of positive responses in the jth cell, and let "$\mathbf{X}_j$" be the vector of predictor variables in the jth cell. The Berkson model postulates the relation

$$P_j = \frac{1}{1 + e^{-\boldsymbol{\beta}\mathbf{x}_j - \xi_j}} \tag{5.2}$$

Note that if $0 < p_j < 1$ for every cell, this is equivalent to

$$\log \frac{P_j}{1 - P_j} = \boldsymbol{\beta}\mathbf{x}_j + \xi_j \tag{5.2a}$$

Linear regression techniques can be applied to equation 5.2*a*. To calculate efficient estimates of $\boldsymbol{\beta}$ requires knowledge of the distribution of ξ. It can be shown

$$\xi_j \sim N(0, n_j\Pi_j(1 - \Pi_j))$$

where n_j = number of observations in the cell and

Π_j = probability of a positive response in the jth cell.

The unknown probabilities (Π_j) can be estimated by the proportions of positive responses (P_j) in each cell and a weighted least squares estimate of β calculated.[3] Because the efficient estimator is only approximated by this procedure, t statistics are true only asymptotically. This WLS logit model was applied to the Leonard data, and the results are reported in the first line of table 5-2.

WLS logit requires more than one observation per cell, preferably many observations per cell, for its reliability depends on accurate estimates of the cell probabilities (Π_j). Furthermore, because the logit transformation is not defined for cells in which $P_j = 0$ or $P_j = 1$, the WLS logit model may require elimination of observations (even if the overall sample size is very large). Both of these considerations severely prejudice results in the example used here. Two of the eight cells had to be deleted, one because $P_j = 0$ and the other because $P_j = 1$.

Berkson (1968, p. 153) suggests an adjustment which precludes the necessity of deleting cells: P_j is replaced by P_j':

$$P_j' = \frac{n_{j,1} + 1/2}{n_j + 1} \tag{5.3}$$

where:

$n_{j,1}$ = number of positive responses in the jth cell and

n_j = the total number of observations in the jth cell

Note that $\log \dfrac{P_j'}{1 + P_j'}$ is defined for virtually every cell; furthermore for $n_j \rightarrow \infty$ P_j' will be extremely close to P_j. WLS logit with the Berkson adjustment was applied to the Leonard data set, and the results are reported in the second line of table 5-2. Inclusion of the two cells which were originally deleted dramatically improves statistics. But the results still suffer from the overall sample size, and are now further prejudiced by application of an admittedly *ad hoc* adjustment. A different approach can be used to avoid these problems.

Maximum Likelihood Logit

Rather than grouping observations into cells, one could define a response function for each individual

Table 5-2
Probability of Adopting Automatic Teller Service
(sample size = 100)

Type of Logit Model Used	*Coefficient on Explanatory Variables, with Asymptotic t Statistics*							
	Intercept	*Age*	*Income*	*Mobility*	*F*	χ^2	R^2	λ
Weighted least squares, empty cells deleted	−0.580 (−0.584)	−0.471 (0.485)	0.805 (0.770)	1.013 (1.102)	0.752		0.530	
Weighted least squares, Berkson adjustment	−0.710 (−0.992)	−0.582 (−0.842)	0.971 (1.350)	1.030 (1.523)	1.734		0.565	
Maximum likelihood	−0.880 (−1.733)[a]	0.822 (−1.621)	1.216 (2.468)[a]	1.129 (2.223)		18.915		0.000

Note: See text for discussion of models and definition of variables; λ is the likelihood ratio for the hypothesis $\beta = 0$ and is a measure of "goodness of fit"; $\chi^2 = -2 \log \lambda$.

[a]Significant at 5 percent level.

$$\Pi_i = F(\beta \mathbf{x}_i) = \frac{1}{1 + e^{-\beta \mathbf{x}_j}} \qquad i = 1, \ldots, n \tag{5.4}$$

where Π_i = true probability that the ith respondent is positive. This relation can be used to define a likelihood function (L) on the observations:

$$L(y_1, \ldots, y_n; x_1, \ldots, x_n) = \prod_{i=1}^{n} F(\beta \mathbf{x}_i)^{y_i} \; [1 - F(\beta \mathbf{x}_i)]^{(1-y_i)} \tag{5.5}$$

where:

y_i = 0, 1 response variable

$\mathbf{x}_i$ = vector of observations on predictor variables

Maximum likelihood estimators (ML logit) can be calculated at relatively low cost, using an iterative computer routine such as the one developed by Nerlove and Press (1973). Asymptotic standard errors in these estimates can be used in calculating asymptotic t statistics. The F statistic of the WLS logit procedure has a counterpart in the likelihood ratio test statistic $= -2 \log \lambda \sim X^2$, where λ = likelihood ratio for $\mathbf{H}_0$: $\beta = 0$.

The likelihood ration (λ) ranges in value from 1 to 0, having smaller values as the goodness of fit improves. It is therefore a possible analog to the coefficient of determination (R^2) in the WLS logit model.

Because the ML logit procedure does not require many observations per cell, what was a "small" sample for weighted least squares may be adequate for maximum likelihood. This feature is evidenced by the dramatic improvement in test statistics with application of the latter (see the third line of table 5-2).

One would expect less difference in results between the WLS logit model and the ML logit model if the sample size were very large. Green, Carmone, and Wachspress (1977) report results of WLS logit estimation of a three-variable, eight-cell model with 10,524 observations. These are replicated in the first line of table 5-3, and test statistics have been calculated in the same manner as was done for the Leonard data set. ML logit estimates for the Green et al. data set are shown in the second line of table 5-3.

Coefficient estimates are very close for both procedures but, interestingly, test statistics are more significant with ML logit, even for this large sample. However, it is clear that difference between WLS logit and ML logit is much related to sample size—minimal for large sample, very pronounced for small samples.

Table 5-3
Probability of Adopting a New Telecommunications Service
(sample size = 10,524)

Type of Logit Model Used	*Coefficients on Explanatory Variables, with Asymptotic t Statistics*							
	Intercept	*Education*	*Income*	*Mobility*	*F*	χ^2	R^2	λ
Weighted least squares	−0.903 (−9.417)[a]	−0.164 (−1.428)	−0.438 (−3.764)[a]	−0.986 (−8.716)[a]	30.884		0.959	
Maximum likelihood	−0.902 (−19.073)[a]	−0.161 (−2.772)[a]	−0.444 (−7.709)[a]	−0.992 (−17.610)[a]		411.526		0.000

Note: Data are from Paul E. Green, Frank J. Carmone, and David P. Wachspress, "On the Analysis of Qualitative Data in Marketing Research," *Journal of Marketing Research* 14 (February 1977):52–59.

[a]Significant at 1 percent level.

Interpretation of Logit Coefficients

One way of presenting results of logit-type models (either WLS or ML) is to calculate predicted probability of a positive response for each cell. These estimates will be superior to the simple proportions in each cell in the sense that they are based on a wider set of information (in logit-type estimation, unlike the simple procedure of calculating cell proportions, sample points in all cells influence estimated probabilities in any one cell).

In addition to predictions for each cell, the researcher may be interested in the marginal effect of changes in each predictor variable. For example, what is the change in probability of adoption implied by increase in income from category 0 to category 1? Define Π^0 as the subject's initial probability of a positive response, and Π^1 as the probability after the change in predictor. Then the induced change in probability is:

$$\Delta\Pi = \Pi^1 - \Pi^0 = \frac{1}{1 + e^{\beta x^1}} - \frac{1}{1 + e^{-\beta x^0}} \tag{5.6}$$

By use of the ML logit estimates for the Leonard data set, these marginal effects are calculated for unit changes in each predictor, for arbitrarily selected initial probabilities of adoption. These figures are reported in table 5-4.

Notice that the magnitude of the marginal effects is influenced in each case by the initial probability of adoption. This indicates a type of interaction among predictor variables, implicit in the logistic function. If one

Table 5-4
Predicted Effects of Changes in Explanatory Variables on Probability of Adopting Automatic Teller Services, Based on Maximum-Likelihood Estimates

	$\Delta\Pi$	$\Delta\Pi$	$\Delta\Pi$
Initial Probability of Adoption = Π	Δ *Age*	Δ *Income*	Δ *Mobility*
.25	−0.123	0.279	0.258
.5	−0.195	0.271	0.256
.75	−0.181	0.160	0.153

Note: Estimates are based on maximum-likelihood estimates of coefficients, reported in table 5-2. See text for discussion of calculation procedure.

assumes for the moment that predictors are continuous, the partial derivates of equation 5.1 are

$$\frac{\partial y}{\partial x_j} = y(1 - y)\beta_j \qquad j = 1, \ldots, k \tag{5.7}$$

Because y is itself a function of all the predictors $(x_1, \ldots, x_k)$, so are the partial derivatives, and this is one meaning of "interaction." This result is generalizable to the case of categorical predictor variables.

Calculations like those in table 5-4 can be especially informative if the number of cells is large. In these cases, rather than selecting arbitrary values for initial probability of a positive response, the researcher might select the quartiles of the predicted probabilities. Marginal effects reported would then be descriptive for the actual sample population.

If the interaction implicit in the logistic function were thought not to capture the full interaction effects, variables could be defined such as:

$$x_{k+1} = \begin{cases} 1 & \text{if } x_1 = 1 \text{ and } x_2 = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$x_{k+2} = \begin{cases} 1 & \text{if } x_1 = 1 \text{ and } x_2 = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\vdots$$

etc.

Inclusion in model 1 of variables like these would alter the partial derivatives in a way which could be said to represent *additional* interaction effects. Note that including all possible interaction effects would be tedious and that the addition of a large number of interaction terms should be weighed against the reduction in the number of degrees of freedom available in estimation.

Suppose that instead of a single dependent variable there were several qualitative dependent variables presumed to be interrelated—e.g., probability of adopting push-button phone service, probability of adopting color phones, and probability of adopting princess style may all be subjects of study and thought interrelated. In such a case a multivariate logistic function is the appropriate mathematical model. To capture interaction effects among the separate but related *dependent* variables it is necessary to include interaction terms (in this instance, bivariate and/or trivariate) indicating coincidence of adoption of the different services. *This* interaction

would not be implicit in separate estimation of marginal logistic functions for each of the three dependent variables unless the extra terms were added.[4]

Conclusions

Many conceptual models can be formulated as probability functions. An efficient and economical technique for estimating such functions is therefore a valuable tool of analysis. The maximum likelihood logit procedure is shown here to be such a technique.

The WLS logit model requires many observations per cell. With survey data, it therefore entails categorization. (With experimental data as in the biological sciences, it need not entail categorization.) The ML logit procedure is based on individual observations and therefore can be used with categorical *or* continuous data.

The WLS logit model requires *ad hoc* adjustment if cells include all positive responses or no positive responses. The ML logit model does not require any such adjustments.

The WLS logit model produces less precise estimates than the ML logit. Sample size may be "too small" for use of the WLS logit model but adequate for ML logit.

Interaction effects are implicit in all logit-type models, and should be taken into consideration in presenting results of these models.

Berkson, the coiner of the term "logit" more than 20 years ago, compared the properties of ML logit and WLS logit estimates (1955, 1968). Using experimental data with many observations per cell, Berkson found little difference between WLS logit (he called it "minimum logit X^2" . . .) and ML logit estimates. But he argued for use of the former on the basis of its relative ease of computation. For marketing research this verdict should now be reversed for two reasons. First, in some marketing applications (unlike experimental-biology applications) there will necessarily be few observations per cell, maybe even only one per cell. Second, the computer revolution has significantly lowered the relative costs of numerical calculations. There is no longer any reason to prefer WLS logit over ML logit for marketing applications.

Notes

1. The maximum likelihood logit model discussed here is slightly different from the model Green et al. identify as a "log-linear maximum likelihood" logit model. The difference between the two maximum

likelihood models is their parameterization. Here a regression format is adopted rather than the ANOVA structure of the Green et al. "log-linear" logit model. The advantages here of the regression-type model are that it (1) facilitates direct comparison with the weighted least squares logit model and (2) highlights an important feature of the maximum likelihood approach: its reliance on individual rather than grouped observations. See Nerlove and Press (1973, pp. 41–45) for an extensive discussion of the relation between the ANOVA (log-linear) maximum likelihood logit model and the regression maximum likelihood logit model.

2. The normal function entails more complex and therefore costly computation. For the maximum likelihood estimator described hereafter, the cost of using the logit formulation was less than one dollar. With probit analysis, the cost of maximum likelihood estimates was greater, but the results were very close.

3. If variances in error terms are not uniform, the OLS estimator is not efficient (meaning that another linear unbiased estimator has less variance). For cases in which residual variances are known (or known up to a scalar constant), a simple transformation of the data (Aitken transformation) restores efficient estimation. This transformation is equivalent to dividing observations by the standard deviations in their respective error terms. Suppose that (in matrix notation)

$$y_j = \mathbf{x}_j'\boldsymbol{\beta} + \xi_j$$

and

$$\text{variance } (\xi_j) = \frac{\sigma^2}{\omega_j}$$

Then the Aitken transformation yields

$$y_j\sqrt{\omega_j} = \mathbf{x}_j'\sqrt{\omega_j}\boldsymbol{\beta} + \xi_j\sqrt{\omega_j}$$

See that residual variances in the equation are uniform:

$$\text{variance } (\xi_j\sqrt{\omega_j}) = \sigma^2$$

When OLS is applied to the weighted data, estimation is efficient.

4. The number of separate logit equations implied by a multivalued criterion is equal, in each case, to the number of independent binary distinctions. If the criterion variable assumes K possible *mutually exclusive* values, the response is completely represented as K-1 binary choices. In the telephone example given in the text, the three phone characteristics are not

mutually exclusive and therefore entail three independent binary choices. In this instance, three separate logit equations would be required. See Nerlove and Press (1973) for a complete discussion of multivariate qualitative models.

References

Berkson, J. (1955), "Maximum Likelihood and Minimum X^2 Estimates of the Logistic Function," *Journal of the American Statistical Association,* 50 (March).

———(1968), "Application of Minimum Logit χ^2 Estimate to a Problem of Grizzle with a Notation on the Problem of 'The Interaction,'" *Biometrics,* 24, 75–96.

Green, Paul E., Frank J. Carmone, and David P. Wachspress (1977), "On the Analysis of Qualitative Data in Marketing Research," *Journal of Marketing Research*, 14 (February):52–9.

Leonard, E.W. (1978), "Life Style Measures and Segmentation in Retail Banking: An Empirical Examination Applied to Automated Tellers," unpublished Ph.D. dissertation, Oklahoma State University.

Nerlove, Marc, and James S. Press (1973), *Univariate and Multivariate Log-Linear and Logistic Models.* Santa Monica: Rand Corporation (December).

6 Utilization of Concept Testing for New-Product Forecasting: Traditional versus Multiattribute Approaches

Edward M. Tauber

A review of multiattribute models by Shocker and Srinivasan (1979) concludes that these approaches for concept evaluation are (potentially) more valuable than traditional concept-testing methods. The purpose of this chapter is to describe and evaluate the use and performance of the traditional versus multiattribute approaches to forecasting the performance of a new concept.

Purpose and Process of Concept Testing

The origin of testing market response to new-product ideas has been traced to Nowland (1947), an industrial designer who, after considerable debate among his peers over the best design for a new flatiron, exposed conceptual renderings to a sample of housewives. While the women rejected every proposed design, the more important discovery was made that indeed new-product concepts could be "tested" before actual production.

Much has been learned since this pioneering experiment. Nevertheless, the purpose and general process of concept testing as reflected in Nowland's effort remain the same today. The general purpose of concept testing is to screen out losers at the idea stage through some form of consumer evaluation. With the ever-rising cost of new-product introductions and the equally high failure rate, management retains a strong interest in this objective. The process of concept testing, regardless of methodology, is basically the same.

Consumers are presented with a *stimulus* (the concept), and *measures* of reaction are taken which the researcher believes are predictive of a *behavioral response,* such as later purchase:

Stimulus	*Surrogate Response*	*Bahavioral Response*
Concept message	Measures of consumer reaction (such as purchase intention or preference)	Purchase of product or service depicted by concept

Traditional and the various multiattribute forms of concept testing differ by how the concept message is presented and by the measures employed to predict purchase. A brief review of these differences and results of studies evaluating each approach will clarify their effectiveness.

Concept Stimulus: Traditional versus Multiattribute

The new-product concept must be communicated to respondents in order to elicit their reaction. Unfortunately, researchers have learned that consumer evaluation of a concept can vary dramatically depending on the nature of the stimulus.

Traditional

When practitioners refer to concept testing, they are talking about a rather standardized procedure employed by most firms (see, for example, McGuire 1973). The basic procedure for concept testing, as practiced today, has its roots in advertising copy testing. The concept, cast in a copy statement or simulated advertisement, is presented to a sample of respondents, and various measurements are taken. Many variables in the presentation of concepts can affect the subjects' responses. Four key influences are:

1. The concept or idea itself
2. The positioning of the idea
3. The form of presentation
4. The execution or design of what is conveyed

Research by Tauber (1972) has revealed that testing just one idea by using different *forms* of communication (for example, a simple printed statement, prototype print advertisement, TV commercial, prototype physical product or package) generates different levels of consumer interest. Haley and Gatty (1971) demonstrated that *execution* of the communication also affects test results. The same concepts executed by different copywriters produced statistically different concept test "scores."

A recent experiment conducted by Queen, Johnson, and Haley (1979) further confirms the differences in result attainable because of confounding of the idea, positioning, form, and execution of the stimulus. Five existing brands each had alternative positionings which were executed in finished advertising. These had been tested for persuasion by using the competitive environment test (CET); for each, one positioning had performed significantly better than the other. Working backward, copywriters wrote a

paragraph concept statement that they believed adequately reflected the positioning inherent in each advertisement. These were then concept-tested by mail with a sample of 500 female household heads reporting purchase intention to each statement. Queen's conclusions were that in all five cases the results were reversed: the concept which scored highest in commercial form scored lowest in the concept test!

One might hypothesize that the reason why consumers responded differently to the advertisement versus the concept statements was strong executional elements. While possible, in one example the bias of different executional elements was eliminated by using the exact same commercial for two completely different positionings. The reversal from commercial to concept test was still true.

In summary, there is much published and unpublished data that support the conclusion that consumer reaction to a concept can be significantly affected by positioning, form of presentation, and execution.

Multiattribute

There are two major types of multiattribute models: the self-explicated (for example, rating importance of attributes) and the full-profile evaluation employed in certain conjoint-analysis approaches.

The most common method of multiattribute concept testing presents respondents with the concept as a *list* of product attributes, consumer benefits, and so on. This approach implicitly assumes that consumers can and will respond to decomposed combination of stimuli (a list of attributes or benefits) in the same way that they will to a "finished" new-product idea. This basic assumption is, as yet, untested.

One reason for a difference is that a bare-bones concept, in attribute list or statement form, communicates *too little* about the idea. When any new idea is tested—a product idea, an advertising idea, a promotion idea—if too little information is provided, respondents can misunderstand it. With an incomplete stimulus, the "need for closure" motivates people to add dimensions (projection), thus resulting in each respondent perceiving the idea differently. Recognizing this danger, many advertising researchers "comprehension-test" new copy. New-product concepts should likewise be comprehension-tested.

Finally, the psychological notion of gestalt argues that people respond to given stimulus in its entirety. It may be unreaslistic to assume that a list of attributes presents the consumer with the same experience generated by a new product in its environment with brand name, packaging, advertising, and similar influences that affect purchase behavior. Traditional concept testing presents respondents with ideas in more of an integrated, gestalt for-

mat (for example, print advertisement). As previously mentioned, research has well documented the variance in consumer response using alternative stimuli.

A newer multiattribute method, conjoint analysis, has the potential for presentation of the stimuli in various forms such as a complete verbal description, advertisement, and so forth. (See Green and Wind 1975.) As such, this method retains the benefit of full presentation of stimuli similar to traditional concept testing.

Concept Measurement: Traditional versus Multiattribute

Regardless of the testing procedure, the measures of consumer reaction to a concept must be predictive of consumer purchase behavior if sales forecasting is the objective. In Nowland's first concept test, he simply asked women if they *liked* the design renderings for a new flatiron. Since that early experiment, a variety of measures have been employed in concept testing.

Traditional

A major advancement in concept testing, deriving from the work of Tobin (1959), Juster (1966), and others, was the incorporation of questions concerning consumer-purchase expectations. Surveys of generalized consumer buying intentions were developed for predicting the demand for durable goods. The Survey Research Center at the University of Michigan, a primary research agency in this field, began this work in 1945 when the Federal Reserve Board sponsored a study of consumer holdings of liquid assets. These studies asked respondents about their intention to buy from various duarable-goods categories in the future, for example, "Will you purchase a car during the next six months?" Purchase-intention scales were adopted in concept testing even though expectations of purchasing abstract new brands may be considerably different from intentions to make a purchase from a generic category such as cars or washing machines. Very few studies have been published testing the predictive validity of purchase-intention scales for nondurables.

Morrison (1979) has attempted to construct a framework for analyzing and interpreting purchase-intention data. Unfortunately, his data base was the purchase-intension data for automobiles and household appliances collected by Juster in 1966. Such data may have little relevance for frequently purchased consumer products. However, his objective of attempting to

understand how purchase intentions link to actual purchase behavior is a meaningful one.

Some empirical work has been conducted in the area of consumer package goods. Axelrod (1968), Haley (1970), Haley and Case (1979), Taylor, Houlahain, and Gabriel (1975), and Clawson (1971) have each published studies on the link of intentions to behavior.

Predictive Validity of Traditional Concept Testing

A validation study conducted by Tauber (1975) revealed some conclusions about purchase intentions. The objective was to determine whether purchase intention asked at concept relates to and can predict the subsequent states or behavior—awareness, trial, and repeat (the elements of a sales forecast). During a five-month time span, six new food products entered test market or went national. Prior to the start of introductory advertising, concept tests were conducted in the test cities, using printed statements derived from the advertising copy. The purchase price was given in the concept statement. Products were disguised in that the actual brand name and corporate name were not used, to avoid predisposing respondents toward the new products that later actually appeared on the market.

After twenty-five weeks following the start of advertising, respondents were recontacted and asked if they were aware, had tried, and had made a repeat purchase of the new products. Thus, by following up the same individuals over time, it was possible to see if they behaved as they said they would in the concept test and, if not, why not.

A cross tabulation of the purchase-intention scale at concept with awareness/unawareness at callback revealed that awareness levels were significantly higher for those who were most interested in the product ideas:

Purchase Intent	*Percentage Aware from Each Group*
Definitely buy	71
Probably	60
Might or might not	54
Probably not	52
Definitely not	38

Apparently, those with interest in the product category pay more attention to commercials for new products in that category, those with no interest more quickly forgot such commercials, or interested parties gain more exposure to the new product in the store since they shop that section.

For cases where respondents were aware of the products, purchase intention differentiated triers from nontriers:

Purchase Intent	*Percentage Tried of Those Aware*
Definitely buy	31
Probably	16
Might or might not	17
Probably not	8
Definitely not	10

Of the methods for analyzing the purchase-intention scale, adding definitely and probably together or taking the top box is incorrect. Definites have a much higher probability of purchase than probablys, and mights are about equal to probablys. A side benefit of this study is the revelation that for trial prediction we actually have a three-point scale: definitely, probably and might or might not, and probably not and definitely not.

While the much higher probability of trial purchase of the definitelys has been emphasized, it should be pointed out that the other categories cannot be ignored. Even though the *odds* of trial from these groups are considerably lower, they tend to be large in numbers. Thus much of *actual* trial comes from these categories. As an example, below is the percentage of total trial coming from each purchase-intention group for the six test-market products:

Purchase Intent	*Percentage of Total Trial from Each Group*
Definitely buy	25
Probably	30
Might or might not	22
Probably not	8
Definitely not	15

For triers of the six test-market products, purchase intent at concept did not differentiate between those who had made repeat purchases and those who had not:

Purchase Intent	*Repeat*
Definitely buy	52%
Probably	43

Purchase Intent	*Repeat*
Might or might not	56
Probably not	50
Definitely not	40

The explanation for why concept test measures are not very useful for predicting repeat is simply that product satisfaction is the critical determinant; this measure, of course, requires trial.

The implication from this validation study is that new-product ideas should not be screened for which ones should be pursued, rejected, or even tested further with *scores* from concept testing since trial is only one element in the sales equation and broad trial is neither necessary nor sufficient for success. An alternative proposal for new-product screening is:

1. Upon first receiving an idea, the new-product manager should estimate all the elements in the sales equation, using judgment and secondary data. The objective of all future research should be to revise and improve estimates of these variables. Those ideas showing the greatest potential sales in the prior estimate should be submitted to the next sequence of consumer testing.
2. After the sales forecast is revised by data from the concept test, any remaining ideas still showing potential should be carried forward. The difference between this procedure and what is often done is that ideas will not necessarily be eliminated if they have low concept scores (because repeat frequency estimates may be high) and products will be ranked by the revised sales forecast, not the "positive purchase-intention score." Likewise, ideas with high purchase-intention scores will not necessarily be pursued if frequency data are low.

The usefulness of concept testing should be to refine the a priori estimate of trial. A second objective of concept testing is the qualitative information provided—target markets are identified with attitudes, demographics, and so on. The most susceptible group can be studied, compared and contrasted with other groups. Advantages, problems, and so forth are also available and important for developing positioning, advertising, and packaging.

Multiattribute Methods

Multiattribute methods employ measures such as "wants," importance ratings, perceptions, preferences, and the like. Traditional concept-testing

measures such as purchase intent are *direct* in asking consumers to forecast their own future behavior. As we have seen, there is a probabilistic relationship of what people say they will do and what they actually do in the future. In contrast, multiattribute methods are *indirect,* and predictions must be inferred from what people say based on their evaluation of the attribute/benefit list. For example, it is assumed that a concept combination which has a higher weighted sum for its attribute values or a greater utility value for a person will be purchased or preferred in place of one with lower evaluations.

Yet, as Srinivasan (1979) has admitted, "Marketing is abundant with examples where two brands may have approximately the same attribute values but enjoy very different market shares." He cites three research examples: Coke and Pepsi have about the same attribute values on sweetness, carbonation, calories, price, and so on, but have quite different shares; two political candidates take approximately the same stand on relevant political issues, enjoy different levels of voter support (for example, the Kennedy's); and several physicians are approximately the same in terms of hours of operation, time to get an appointment, waiting and travel time, but have different shares.

Srinivasan's hypothesized reasons for this phenomenon include the following: (1) relevant attributes are excluded; (2) brands may differ in marketing pressure (distribution, advertising, and so forth); (3) imagery; and (4) brands introduced earlier perform better in the market than "me too" brands.

This is quite reasonable. However, if multiattribute measurements can be inaccurate for such reasons when modeling existing brands and markets, these problems are heightened considerably for new-product concepts. Unfortunately, marketers are rarely sure of all the relevant dimensions that will determine consumer interest in a new product. The many new-product failures and multiple reasons for failure are testimony to this uncertainty. Estimating the impact of imagery of a *new* brand is exceedingly difficult at the concept stage since image effects (to the extent they will exist) tend to build over time—the Charlie girl, Pepsi generation, and the like. Of course, the bare-bones attribute list minimizes measuring any imagery.

Finally, most multiattribute models implicitly assume the contribution of each attribute or benefit dimension to individual choice to be independent of the others. In product-formula optimization, interaction effects are often as important as main (independent) effects. The assumption of orthogonality may be the most tenuous assumption when these models are used in the new-product area. An early theory by Tauber (1972) was that successful new-product ideas are the result of synergy for certain combinations of elements. Psychologists concur that new ideas are a "combination of concepts" already in the mind. Creative innovations are the result of the synergy of two or more elements that alone have only average merit. For example successful Stove Top Stuffing Mix is the combination of existing

concepts of breaded stuffing and mix; Cup of Soup is instant soup and a cup; and Pop Tarts are the elements of pastry and toaster. Summing independent attribute ratings is conceptually diametrically opposed to the theories of creativity in psychology.

Some empirical evidence exists for the predictive validity of multiattribute methods and conjoint analysis in particular. An excellent review of cases by Green and Srinivasan (1978) reveals a positive track record. However, a number of the cases are not analyses of new product evaluations. Thus, it remains for future work to provide stronger evidence in the new product testing area.

Utilization of Multiattribute Methods of Concept Testing for Sales Forecasting

The newer multiattribute methods have been employed selectively in attempting to identify the structure of a given market. In this area they have proved quite useful. However, adoption of these model-based systems for sales forecasting has been limited. As with any "new product," adoption and frequency of use are a function of product satisfaction. As mentioned, it may be that it is not possible to start with a combination of attributes and expect it to represent the new idea.

Another reason for the poor adoption of multiattribute methods for forecasting is the recognition by most practitioners that new products which achieve immediate success are generally three degrees away from existing products, not one hundred eighty degrees. What this implies is that in new-product generation and screening, evolutionary changes are more likely to attain success than revolutionary ones which require a much longer time for consumer and trade acceptance. The inherent *complexity* built into a number of the multiattribute methods assumes the new big idea to be a complex variation from what presently exists. In reality, major breakthroughs in a product category generally result from technological advances, not complex marketing repositioning.

For the near future, the multiattribute methods will remain in the realm of experimentation until they have demonstrated that they are accurate forecasters of consumer response to a concept. Traditional concept-testing methodology will continue to dominate practice.

References

Axelrod, J.N. "Attitude Measures that Predict Purchase," *Journal of Advertising Reseach,* March 1968, pp. 3-18.

Clawson, C. Joseph. "How Useful are 90 Day Purchase Probabilities?" *Journal of Marketing,* October 1971, pp. 43-47.

Green, Paul E., and Srinivasan, V. "Conjoint Analysis in Consumer Research: Issues and Outlook." *Journal of Consumer Research* 5(1978):103–123.

Green, Paul E., and Wind, Yoram. "New Way to Measure Consumers' Judgments." *Harvard Business Review,* July–August 1975, p. 107.

Haley, Russell. "We Shot an Arrowhead (#9) into the Air." *Proceedings.* Advertising Research Foundation, 1970.

Haley, Russell, and Case, Peter B. "Testing Thirteen Attitude Scales for Agreement and Brand Discrimination." *Journal of Marketing* 43(Fall 1979):20–32.

Haley, Russell, and Gatty, Ronald. "The Trouble with Concept Testing." *Journal of Marketing Research* 8, no. 2(1971):23.

Juster, Thomas F. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." *Journal of the American Statistical Association,* no. 61(September 1966):658–696.

McGuire, E. Patrick. *Evaluating New Product Proposals,* no. 604. New York: The Conference Board, 1973.

Morrison, Donald G. "Purchase Intentions and Purchase Behavior." *Journal of Marketing* 43(Spring 1979):65–74.

Nowland, Roger L. "Pre-Design Research as Applied to Product Development." *American Society of Mechanical Engineers,* paper no. 47–A–57, December 1947.

Queen, Penelope; Johnson, Tod; and Haley, Russell. "An Experiment in Concept Testing for Existing Brands for Dancer Fitzgerald Sample." Unpublished study, 1979.

Shocker, Allan D., and Srinivasan, V. "Multi-Attribute Approaches for Product Concept Evaluation and Generation: A Critical Review." *Journal of Marketing Research,* no. 16 (May 1979):159–180.

Srinivasan, V. "Network Models for Estimating Brand-Specific Effects in Multi-Attribute Marketing Models." *Management Science* 25, no. 1(January 1979):11–21.

Tauber, Edward M. "HIT: Heuristic Ideation Technique, A Systematic Procedure for New Product Search." *Journal of Marketing,* January 1972.

———. "What Is Measured by Concept Testing?" *Journal of Advertising Research* 12(December 1972):35.

———. "Predictive Validity in Consumer Research." *Journal of Advertising Research,* October 1975, pp. 59–64.

Taylor, James W.; Houlahain, John J.; and Gabriel, Alan C. "The Purchase Intention Question in New Product Development: A Field Test." *Journal of Marketing* 39(January 1975):90–92.

Tobin, James. "On the Predictive Value of Consumer Intentions and Attitudes." *Review of Economics and Statistics,* no. 41(February 1959):1–11.

Part III Pre-Test-Market-Based New-Product Forecasting Models

Test market, the final stage in the new-product development process, is often considered a necessary safeguard against the introduction of "wrong" new products. Many companies require test market as a mandatory stage prior to the final approval of the national-market introduction of a new product. Yet, more than any other stage in the new-product development, alternatives to test market are often sought. This search for alternative approaches is motivated by the high cost and other limitations of test market.[1]

Some of the suggested alternatives to test market include in-home use tests, central-location (laboratory simulation) tests, mini test markets, regional rollouts, and adaptive experimentation.[2] Whereas all the five suggested procedures can be considered as alternatives to test market, the first three can also be used as a precursor stage to test market. That is, given the high direct and indirect cost of test market and its other limitations, one may employ any of the first three procedures to help determine whether to go or not go with test market.

In addition to the review (chapter 7) and industry-utilization (chapter 11) chapters, this part contains three chapters (8, 9, 10) dealing with "laboratory" pre-test-market evaluation of new products (that is, central-location tests). The primary objective of such pre-test-market evaluation procedures is to simulate the awareness-trial-repeat-purchase process using controlled laboratory and product-use tests. The specific procedures included are ASSESSOR, developed by Management Decision Systems, Inc., and described in chapter 8; LTM estimating procedures, offered by Yankelovich, Skelly, and White, Inc., in chapter 9; and the COMP system, offered by Elrick and Lavidge, Inc., and described in chapter 10.

In chapter 7 Robinson compares and contrasts these three approaches with several other market-measurement procedures including the one offered by his own firm (that is, SPEEDMARK) and reference procedures such as conventional test market. His comparison is based on twelve major and thirty-seven detailed characteristics which are presented and clustered, highlighting the similarities and differences between the procedures.

Finally, the chapter on industry utilization by Levine (chapter 11) argues that the pre-test-market research is still emerging as a tool that management can use and many do, but the procedures have not yet been fully validated and the need for this kind of information is not fully

explained or developed. He describes the uses that pre-test-marketing research can be put to, outlines the limitations and advantages of pre-test-market research, and indicates areas for further development.

Notes

1. For a discussion of the limitations of test-market procedures, see Yoram Wind, *Product Policy: Concepts, Methods, and Strategy* (Reading, Mass.: Addison-Wesley, 1981). chap. 14.
2. For a discussion of these five approaches, see Wind, Ibid., ch. 14.

7 Comparison of Pre-Test-Market New-Product Forecasting Models

Patrick J. Robinson

The measurement of total market potential for a new product presents a great challenge to the modern manager. Curiously, perhaps the greatest hurdle to scientific innovation in marketing is the time-honored tradition of test market and limited rollouts as practiced for many decades by many consumer packaged-goods marketers. In these traditional approaches, the marketer produces enough of a product to place it on the market in the belief that these pilot test results can give a realistic preview of how a product would fare in actual full-scale rollout. The principal virtue of such conventional market-test centers on their intuitive appeal to "practical" brand and product managers—and other administrators who find it easy to relate to "results" from the "real world." Yet, all too often, these pragmatic test procedures fail miserably or provide inconclusive or misleading evidence. They are "goldfish-bowl laboratories" in which many unplanned forces may enter and many unplanned and uninstrumented changes can occur to obscure and confound the typical measures of success versus failure of a new or restaged product launch.

In contrast to traditional test markets and limited rollouts—both of which require substantal money, time, and resource commitments—there has been a growing interest in, and appreciation for, the modern tools of survey research and system simulation. These computer-assisted representations of a market microcosm offer substantial management insights to the well-informed user. Such simulations can also pose potential pitfalls for the naive or uniformed user who does not become sufficiently involved in considering both the explicit and the *implicit* assumption structures on which the simulations are parameterized and tested for sensitivity.

In recent years a number of commercial pre-test-market measurement procedures have been developed and implemented. This chapter compares some of these procedures, including PURCHASE ACTION (ASI Market Research, Inc., New York), ASSESSOR (Management Decision Sciences, Inc., Weston, Mass.), COMP (Elrick & Lavidge, Chicago), LTM (Yankelovich, Skelly & White, New York), SPEEDMARK® (Robinson Associates, Bryn Mawr, Penn.), and MICROMARKET (Tele-Research, Los Angeles).

These six procedures are compared among themselves and contrasted with traditional test-market approaches and conventional marketing rollouts as, de facto, other practical (albeit costly and time-consuming) market-measurement procedures. Thirty-seven detailed measurement attributes are used in the comparison of these procedures. Judgmental ratings were assigned (by the author, based on public-domain data and informed opinions of professional colleagues) to the eight market-measurement procedures employing all thirty-seven basic attributes. The data sources used for this evaluation are listed under "References" at the end of this chapter. Since these value judgments are based on the available public information concerning the known options, then, to the extent that discrepancies or oversights exist or that modifications may be made in these competing approaches, some of what follows deserves interpretive refinement by the reader. At the time of this writing, such modifications are thought to be relatively minor insofar as the comparisons being made here. [Since the completion of this analysis, a new simulated test-market service LITMUS has come to our attention as an outgrowth of LTM. See Blackburn and Clancy (1980). Also, system of concept and product testing, pre-test-market measurement, and test-market tracking called BASES has been introduced by the Burke Companies.]

The thirty-seven measurement attributes, grouped into twelve aggregate characteristic sets, are briefly described below, along with a conceptually idealized position for each attribute as an idealized benchmark or standard of comparison.

Scope of Test-Product Configuration

Any pre-test-market forecasting model should test the correct set of product configurations, a concept which includes the packaging and pricing as well as the product formulation itself. Ideally, all product configurations which are likely to be introduced into the marketplace should be encompassed in the design and scope of the pre-test-market forecasting model.

Scope of Test-Marketing Plan

Similarly, the five marketing-plan components—advertising budget, promotional expenditures, media mix, deal pricing, and sampling—should be included in any ideal procedure, and have been included as basic attributes in this framework. Other marketing-plan variables may arise in special situations and could dictate suitable modifications in test procedures and marketplace implementations.

Scope of Test-Environment Conditions

In the marketplace, external factors invariably affect and perturb the performance of any new product. The most important external factors which affect most or all product categories are:

1. Competitive product positioning, relative weight and mix of promotional activities (notably advertising and sales promotion)
2. The set of competitive products available through distribution in the marketplace
3. Buyers' benefit-seeking propensities, personal values, attitudes, and use in the category and other buying influences including reference-group pressures and confidence in the state of the economy

These three variables have been selected as basic attributes which optimally would be testable at all feasible levels by a pre-test-market forecasting model. Other external factors may be important in a particular market.

Measurement of Buyer Purchase Behavior

The measurements taken in an optimal or idealized pre-test-market forecasting model would be the natural pattern of product purchases among a realistic set of brand choices. The three basic system simulation and operational gaming attributes which can achieve the necessary and sufficient verisimilitude for these aggregate characteristics are a fully realistic set of brand choices, the opportunity to purchase the product at any point in time, and the absence of artificial constraints on test respondents' purchasing power. (Environmental constraints on purchases are considered as a separate characteristic.) The optimal constraint on respondents' purchasing power, in terms of realism, would be the respondent's own household budget or discretionary income.

Conditions of Environmental Exposure—Test Products and Promotion

The environment in which test respondents are exposed to products and promotion should be realistic. The setting of product purchases (that is, the actual or simulated store or in-home task) should be as close as possible to the respondent's store of choice in respect to ambiance and evoked buyer behavior. The second basic attribute in this category is the advertising context, which should be as close as possible to the media ordinarily seen by the respondent. A third basic attribute, which is applicable when sales promo-

tion such as sampling is used, is the sampling channel of distribution. If sampling by mail is part of the new product's anticipated marketing plan, then sampling by mail should be used in testing. The environmental realism of other promotional programs has not been explicitly included in the set of basic attributes. However, the advertising and sampling environment used in the test procedure can be a reasonable proxy of the overall promotional environment used.

Environmental Exposure—Competitive Products and Promotion

This characteristic is related to the scope of the test described above, in the sense that a realistic choice set of competitive products is required for the forecast. Here, however, it is the purchase environment rather than the forecasting model that is required to include a realistic competitive brand set. The same is true in regard to the need for regularly seen competitive advertising by test respondents. The purchase environment of each respondent should include normally seen competitive advertising.

Conditions of Measurement

The previous two characteristics dealt with the environmental exposure of test respondents, while the characteristic listed before these specified which measurements should be taken. The environment in which these measurements are taken is important and includes four basic attributes: first, the unobtrusive instrument or measurement procedure (for example, the absence of a "Hawthorne effect" or instrument error); second, the frequency of measurement (as distinct from the frequency of purchase opportunity already mentioned); third, the incorporation of alternatives into the measurement procedure; and fourth, the degree of proximity between the generation of purchase data and their recording. Ideally, the measurement environment should also include measurements under alternative circumstances. These four attributes capture the essential ingredients of an ideal measurement environment.

Sampling Procedures

Once the issues of what should be measured and under what circumstances are addressed, it remains to address the issue of selecting subjects for such measurement. Four basic attributes of the sampling procedure are considered here. First, what should the sample element be? Ideally, it should be the unit which is normally responsible for making the purchase decision.

(This will usually be an appropriate individual, but it could be a family or, in industrial situations, a more complex buying unit.) The second basic attribute here is the sample universe from which sample elements are drawn. Ideally, the sample universe will be inclusive of all purchase units of interest, with no restrictions. The third basic attribute involves the method of drawing the sample, which should be probabilistic, thus allowing inferences to be drawn scientifically from the sample to project onto the population. The fourth basic attribute concerns sample size, which should be sufficiently large to allow statistically sound conclusions to be arrived at, relatively free of the "luck of the draw." Restated, the forecasts generally should be theoretically reproducible if a different (sufficiently large) sample were to be drawn. These four basic attributes summarize the essential sample-plan qualities that are needed by a pre-test-market forecasting model.

Projectability of Results

An aggregate characteristic of obvious importance is the end result of the model, namely the information provided to the forecast user. There are two basic attributes of the projectability or range of applicability of the results: the time horizon or endpoint of the forecast and the ability of the model to provide forecasts or demand patterns for any prespecified time period within that horizon.

Assumptions Required

No forecasting model can be better than the assumption structure which underlies it. Two distinct basic attributes can be identified. The first is the reasonableness of the assumptions utilized in the model. The degree of reasonableness of an assumption can be thought of as the extent to which an informed panel of appropriate experts would agree as to its validity. The second distinct aspect of the assumption structure is the extent to which the forecasts are dependent on key explicit or implicit assumptions. "Robust" techniques which produce valid results independently of the specific assumptions used are desirable.

Model Specification

The heart of any pre-test-market forecasting model is the set of mathematical equations, and the resulting computer programs, which convert input data to forecast outputs. The set of equations may include algorithms, heuristics, and other judgmental factors or parameters and boundary constraints.

Three basic modeling attributes have been identified. The first is the completeness of the model specification; that is, are all relevant variables accounted for in the model? Variables which are not explicitly included will be defined by imputation or default, and are therefore likely to be ignored by the analyst. A second attribute is the correctness of the model specification for those variables which the model incorporates. This attribute is probably the most difficult to assess. The third attribute in this set is the decomposition of the model, that is, the degree to which the model can be analyzed and evaluated with respect to the correctness of its specification, internal consistency of units of analysis (measures), completeness of assumptions and constraints, and appropriateness of its other attributes. Clearly, an undefinable or unspecified black-box model is of questionable or doubtful utility.

Operational Characteristics

Three operational characteristics of these procedures have been included: time required for completion, approximate range of costs, and competitive readings. The importance of all three attributes is obvious. Other operational characteristics are highly relevant, but can best be assessed in relation to an individual user's requirements. Characteristics which have been excluded from the present analysis include amount of experience in the user's product category, degree of success or failure in previous forecasting efforts, competence and character of individuals responsible for project execution and support, reliability and stability of firm which offers services, and flexibility in adapting to requirements of individual users' testimonials and disclaimers by the supplying firms or their clients. Of course, these operational characteristics can often be the decisive factors in selecting a pre-test-market forecasting service.

The thirty-seven disaggregate measurement attributes (and their twelve collective attribute sets) provide a uniform basis for understanding the structure and comparing the various pre-test-market forecasting models. Table 7–1 presents a descriptive comparison among all the previously listed services, along with conventional test markets and limited rollouts. This comparison chart presents a verbal summary of each feature, along with a numerical rating on a scale of 1 to 5, with 5 being the best or optimal or ideal level (also listed as a benchmark or par) and 1 being the least desirable or most inappropriate level.

Since the measurement and policy implications of the resulting fairly substantial data matrix are not readily apparent on inspection, these comparisons can be facilitated through the use of numerical taxonomy analysis to provide a performance-space portrayal of the eight measurement tech-

Table 7–1
Comparison Chart for Pretest-Market Forecasting Models
(ratings in lower, left-hand corner are on a 1-to-5 scale)

General Characteristics	*Basic Attribute*	*Optimal Level of Attribute (1)*	*Limited Rollout (2)*	*Conventional Test Market (3)*
Scope of Test-Product Configuration	Product formulations tested	All feasible product formulations 5	One product only 1	One product only 1
	Packaging alternatives	All feasible packaging alternatives 5	One package only 1	Two or three packages testable using split sample 2
	Pricing alternatives	All feasible pricing alternatives 5	One price only 1	Two or three price levels testable using split sample 2
Scope of Test-Marketing plan	Advertising budgets tested	Range of advertising 5	Single advertising budget 1	Different advertising budgets possible for different test markets 3
	Distribution of promo-expenditures	Variable distribution of expenditures 5	Rollout distribution only 1	Effects of different distributors possibly confounded with total budget differences 2
	Media mix	Variable media mixes 5	Single media mix 1	Different media may be used in different test markets 3
	Deal pricing	Includes effects of various deal/coupon pricing 5	Usually includes one or more deals where appropriate 3	Limited use of deal pricing 2
	Sampling	Includes effects of feasible sampling programs 5	Usually includes sampling where appropriate 3	Limited use of sampling 2

Table 7–1 continued

General Characteristics	*Optimal Level of Attribute*	*Optimal Level of Attribute (1)*	*Limited Rollout (2)*	*Conventional Test Market (3)*
Scope of Test-Environment Conditions	Competitive marketing activity	All possible competitive promotion 5	Single but realistic competitive environment 3	Test-market "sabotage" possible 2
	Competitive product formulations tested	All possible competitive product formulations 5	Actual single set of competitive products/prices/packages 3	Test-market "sabotage" (e.g., price changes possible) 2
	Economic conditions	Feasible range of economic conditions in forecast time horizon 5	Actual conditions throughout U.S. 4	Actual conditions in test-market areas only 3
Measurement of Buyer Purchase Behavior	Number of purchase opportunities	Opportunity to purchase at any time 5	Purchase opportunity at any point within rollout timetable 5	Purchase opportunity at any within test-market timetable 4
	Constraints on purchasing power	Purchases limited only by household budgetary constraints 5	Purchases limited only by household budgetary constraints 5	Purchases limited only by household budgetary constraints 5
	Realism of purchase choices	Fully realistic domain of choices 5	All product available to consumer during test period 5	All product available to consumer during test period 5

Conditions of Exposure—Environment	Purchase setting	Purchases made in store of choice 5	Purchases made in store of choice 5	Purchases made in store of choice 5
	Advertising setting	Advertising in media normally seen 5	Advertising in media normally seen 5	Media selection restricted to those in test-market areas 4
	Sampling channels	Sampling done using normal channels 5	Sampling done using normal channels 5	Sampling channels restricted to those available in test-market area 4
Competitive Context	Competitive product availability	Competitive product set found in store of choice 5	Competitive product set found in store of choice 5	Competitive product set found in store of choice 5
	Competitive promotional activity	Competitive advertising which is regularly seen 5	Competitive advertising which is regularly seen 5	Possible test-market "sabotage" by competitors 3

Table 7–1 continued

General Characteristics	*Basic Attribute*	*ASSESSOR (4)*	*SPEEDMARK® (5)*	*COMP (6)*
Scope of Test-Product Configuration	Product formulations tested	Two or three products testable using split sample 2	Thousands of alternatives testable using conjoint scaling 4	Two or three products testable using split sample 2
	Packaging alternatives	Two or three packages testable using split sample 2	Variety of packaging alternatives tested through conjoint scaling 4	Two or three packages testable using split sample 2
	Pricing alternatives	Two or three price levels testable using split sample 2	Price-elasticity curve estimated for segmented market 4	Two or three price levels testable using split sample 2
Scope of Test-Marketing Plan	Advertising budgets tested	Model incorporates assumed sales = advertising response curve 3	Model incorporates assumed sales = advertising response curve 3	Model incorporates assumed sales = advertising response curve 2
	Distribution of promotional expenditures	Distribution not explicitly dealt with 2	Model incorporates effects of distribution 3	Distribution not explicitly dealt with 3
	Media mix	Partial assessment of effects of media mix 3	Partial assessment of effects of media mix 3	Partial assessment of effects of media mix 3
	Deal pricing	Deal pricing not explicitly dealt with 2	Model contains deal price-elasticity variable 4	Deal pricing not explicitly dealt with 2
	Sampling	Sampling tested using nonpurchasers of test products 4	Sampling tested using mail or door-to-door distribution 4	Sampling tested using mail or door-to-door distribution 3
Scope of Test-Environment Conditions	Competitive marketing activity	Competition not explicitly dealt with 2	Effects of competitive promotion in model 4	Competition not explicitly dealt with 2

	Competitive product formulations tested	Competition not explicitly dealt with 2	Effects of competitive products in model 3	Competition not explicitly dealt with 2
	Economic conditions	Economic conditions not explicitly dealt with 2	Economic conditions not explicitly dealt with 2	Economic conditions not explicitly dealt with 2
Measurement of Buyer Purchase Behavior	Number of purchase opportunities	Trial purchase and single repeat purchase 2	Trial purchase and up to three repeat purchases 3	Trial purchase and repurchase intention 2
	Constraints on purchasing power	Respondent given money or coupons sufficient to purchase item 3	Respondent given money or coupons sufficient to purchase two items 3	Respondent given money or coupons sufficient to purchase item 3
	Realism of purchase choices	Display of most widely used current available products 3	Display of most widely used current available products 3	Display of most widely used current available products 3
Conditions of Exposure—Environment	Purchase setting	Purchases made in laboratory setting 2	Purchases made from in-home mock shelf display or in-store display 2	Purchases made in laboratory setting 2
	Advertising setting	Respondents sees test product and competitor's commercials in sequence 3	Respondent sees test product and competitor's commercials in sequence 3	Respondent sees test product and competitor's commercials in sequence 3
	Sampling channels	Samples given to respondents who do not purchase test product 2	Sampling by mail or door-to-door prior to depth interview 3	Samples given to respondents who do not purchase test product 2
Competitive Context	Competitive product availability	Selected set of products with overall widest use 4	Selected set of products with overall widest use 4	Selected set of products with overall widest use 4
	Competitive promotional activity	Respondents sees product-category commercials and print advertising 4	Respondent sees product-category commercials and print advertising 4	Respondent sees product category commercials and print advertising 4

Table 7-1 continued

General Characteristics	*Basic Attribute*	*LTM* *(7)*	*TELE-RESEARCH* *(8)*	*ASI* *(9)*
Scope of Test-Product Configuration	Product formulations tested	Two or three products testable using split sample 2	Two or three products testable using split sample 2	Two or three products testable using split sample 2
	Packaging alternatives	Two or three packages testable using split sample 2	Two or three packages testable using split sample 2	Two or three packages testable using split sample 2
	Pricing alternatives	Two or three price levels testable using split sample 2	Two or three price levels testable using split sample 2	Two or three price levels testable using split sample 2
Scope of Test-Marketing Marketing Plan	Advertising budgets tested	Judgmental assessment of effects of advertising budget 2	Judgmental assessment of effects of advertising budget 2	Judgmental assessment of effects of advertising budget 2
	Distribution of promotional expenditures	Distribution not explicitly dealt with 2	Distribution not explicitly dealt with 2	Distribution not explicitly dealt with 2
	Media mix	Media mix not explicitly dealt with 2	Media mix not explicitly dealt with 2	Media mix not explicitly dealt with 2
	Deal pricing	Deal pricing not included 1	Deal pricing not included 1	Deal pricing not included 1
	Sampling	Sampling not included 1	Sampling not included 1	Sampling tested using non-purchases of test 3
Scope of Test-Environment Conditions	Competitive marketing activity	Competition not explicitly dealt with 2	Competition not explicitly dealt with 2	Competition not explicitly dealt with 2
	Competitive product formulations tested	Competition not explicitly dealt with 2	Competition not explicitly dealt with 2	Competition not explicitly dealt with 2

	Economic conditions	Economic conditions not explicitly dealt with 2	Economic conditions not explicitly dealt with 2	Economic conditions not explicitly dealt with 2
Measurement of Buyer Purchase Behavior	Number of purchase opportunities	Trial purchase and repurchase intention 2	Trial purchase and repurchase intention 2	Trial purchase and repurchase intention 2
	Constraints on purchasing power	Respondent given partial payment only, as seed money 2	Respondent given partial payment only, as coupon 1	Respondent given money or coupons sufficient to purchase item 3
	Realism of purchase choices	Display of most widely used current available products 3	Display of most widely used current available products 3	Respondents must purchase product in unrelated category as well 1
Conditions of Exposure—Environment	Purchase setting	Purchases made in laboratory setting 2	Purchases made in actual store 4	Purchases made in laboratory setting 2
	Advertising setting	Respondent sees test product and competitor's commercials in sequence 3	Respondent sees commercials in unrelated as well as test-product category 4	Respondent sees commercials in unrelated as well as test-product category 4
	Sampling channels	Sampling not included in test 1	Sampling not included in test 1	Samples given to respondents who do not purchase test product 2
Competitive Context	Competitive product availability	Selected set of products with overall widest use 4	Product set found in test store 5	Selected set of products with overall widest use 4
	Competitive promotional activity	Respondent sees product-category commercials and print advertising 5	Test product category not identified 5	Respondent sees commercials in two product categories 4

Table 7–1 continued

General Characteristics	*Basic Attribute*	*Optimal Level of Attributes (1)*	*Limited Rollout (2)*	*Conventional Test Market (3)*
Conditions of Measurement	Intrusiveness of measurement procedures	Unobserved nonintrusive 5	Unobserved nonintrusive 5	Unobserved nonintrusive 5
	Frequency of measurement	Continuous audit throughout time horizon 5	Continuous audit during testing period 4	Continuous audit during testing period 4
	Alternatives incorporated in measurement	Measurements include feasible set of alternatives 5	Single set of conditions 1	Possible differences among test-market area 2
	Measurement environment	Measurements taken at point of data generation 5	Measurement using store audit 4	Measurement using store audit 4
Sampling procedures	Sample element	Individual purchase unit 5	Individual purchase unit 5	Individual purchase unit 5
	Sample restrictions	Unrestricted sample 5	Unrestricted sample 5	Sample restricted to test-market areas 2
	Sample methodology	Pure probability sample 5	Pure probability sample 5	Sample restriced to test-market areas 2
	Sample size	Sufficiently large sample to replicate results 5	Sufficiently large sample to replicate results 5	Large sample in test-market areas 4
Projectibility of Results	Projectibility in specified time interval	Results projectible for any specified time interval 5	Projectible to any interval within testing period 4	Limited by representativeness of test marketing 2

	Duration of projectibility	Results projectible to end of planning horizon 5	Projectible to end of testing period 3	Limited by representativeness of test marketing 2
Assumptions Used	Degree to which assumptions influence results	Results are valid independent of assumptions made 5	Results are valid independent of assumptions made 5	Assumption that test markets represent total market 4
	Reasonableness of assumptions	Assumptions universally accepted 5	No assumptions needed 5	Assumption that test markets represent total market 3
Model Specification	Correctness of specification for included variables	Included variables are correctly specified 5	No variables needed 5	No variables needed 5
	Completeness of specification	All relevant variables included in model 5	Alternative product/promotion not included 2	Alternative product/promotion not included 2
	Dissection of Model	Model fully selfcontained and analyzable 5	Effects of marketing variables are confounded 3	Effects of marketing variables are confounded; also effects of sabotage and test areas 2
Operational Characteristics	Time required	Test completed in time to preempt competition and maximize impact	Up to 1 year	Usually 3 to 6 months
	Cost	Cost insignificant relative benefits	$2 to $10M	$0.5 to $1.5M
	Competitive readings	Secrecy ensured 5	Competitive activity integral part of rollout 1	Competitive readings and/or sabotage possible 2

Table 7–1 continued

General Characteristics	*Basic Attribute*	*ASSESSOR (4)*	*SPEEDMARK® (5)*	*COMP (6)*
Conditions of Measurement	Intrusiveness of measure-procedures	Telephone and depth interview plus respondent purchase 3	Telephone and depth interview plus respondent purchase 3	Depth interview plus respondent purchase 3
	Frequency of measurement	Trial plus one repeat purchase 2	Trial plus one to three repeat purchases 3	Trial plus one repeat purchase 2
	Alternatives incorporated in measurement	Model allows testing of some alternatives 3	Alternatives tested simultaneously using conjoint scaling 4	Model allows testing of some alternatives 3
	Measurement environment	Measurements taken at times of purchase opportunities 3	Measurements taken at times of purchase opportunities 3	Measurements taken at times of purchase opportunities 3
Sampling procedures	Sample element	Single buying individual 4	Single buying individual 4	Single buying individual 4
	Sample restrictions	Sample restricted to specific mall shoppers 2	Restriction to current users of product category 4	Sample restricted to specific mall shoppers 2
	Sample Methodology	Convenience sample of mall shoppers 2	Cluster probability sample 4	Convenience sample of mall shoppers 2
	Sample size	300 respondents more if split samples required 4	300 respondents more if split samples required 4	300 respondents more if split samples required 4

Projectibility of Results	Projectibility in specified time interval	Steady-state projections only 1	Dynamic model used to project for monthly intervals 4	Steady-state projections only 1
	Duration of projectibility	Steady-state projections only 3	Model projects sales up to 2 years postintroduction 3	Steady-state projections only 3
Assumptions Used	Degree to which assumptions influence results	Market forecast model uses assumptions which vary in influence on results 3	Market forecast model uses assumptions which vary in influence on results 3	Market forecast model uses assumptions which vary in influence on results 3
	Reasonableness of assumptions	Generally reasonable or accepted assumptions 3	Generally reasonable or accepted assumptions 3	Generally reasonable or accepted assumptions 3
Model Specifications	Correctness of specifications for included variables	Specification based on published market models 4	Specification based on published market models 4	Specification based on published market models 4
	Completeness of specification	Model includes interactions among important competing brands 4	Also incorporates some alternatives 4	Model includes interactions among important competing brands 4
	Dissection of model	Model widely published but contains judgmental reconciliation of two predictions 3	Model substantially free of subjective factors; but critical review has been limited 4	Model documented in working papers, substantially free of subjective factors 4
Operational Characteristics	Time required	1 to 3 months	1 to 3 months	1 to 3 months
	Cost	\$25 to \$50K	\$35 to 50K	\$25 to 50K
	Competitive readings	Competitive readings highly unlikely 5	Competitive readings highly unlikely 5	Competitive readings highly unlikely 5

Table 7–1 continued

General Characteristics	*Basic Attribute*	*LTM (7)*	*TELE-RESEARCH (8)*	*ASI (9)*
Conditions of Measurement	Intrusiveness of measurement procedures	Respondents subject to influences of their groups 2	Telephone and depth interview plus respondent purchase 3	Telephone and depth interview plus respondent purchase 3
	Frequency of measurement	Trial plus one repeat purchase 2	Trial purchase only 1	Trial purchase only 1
	Alternatives incorporated in measurement	Alternatives possible using split sample 2	Alternatives possible using split sample 2	Alternatives possible using split sample 2
	Measurement environment	Group interview used after purchase 2	Measurements taken at time of purchase opportunities 3	Measurements taken at time of purchase opportunities 3
Sampling Procedures	Sample element	Respondents subject to group influences 3	Single buying individual 4	Single buying individual 4
	Sample restrictions	Sample restricted to specific mall shoppers 2	Sample restricted to specific shoppers 2	Sample restricted to specific mall shoppers 2
	Sample methodology	Convenience sample of mall shoppers 2	Convenience sample of shoppers 2	Convenience sample of mall shoppers 2

	Sample size	300 respondents more if split samples required 4	300 respondents more if split samples required 4	300 respondents more if split samples required 4
Projectibility of Results	Projectibility in specified time interval	Steady-state projections only 1	No market-share estimates 1	No market-share estimates 1
	Duration of projectibility	Steady-state projections only 3	Steady-state projections only 1	Steady-state projections only 1
Assumptions Used	Degree to which assumptions influence results	Results depend strongly on assumptions made 2	Results are free of assumptions 5	Results are free of assumptions 5
	Reasonableness of assumptions	Generally reasonable or accepted assumptions 3	Conversion of test data to market-share forecasts difficult 2	Conversion of test-data to market-share forecasts difficult 2
Model Specification	Correctness of specification for included variables	Specification sparsely documented 3	Norms provide aid in interpreting results 2	Norms provide aid in interpreting results 2
	Completeness of specification	Specification sparsely documented 3	Alternative products/promotion not included 2	Alternative products/promotion not included 2
	Dissection of Model	Specification sparsely documented 2	No model specification 2	Model fully self-contained and analyzable 5
Operational Characteristics	Time required	1 to 3 months	2 to 4 weeks	2 to 4 weeks
	Cost	$20 to $30K	$5 to $8K	$8 to $12K
	Competitive readings	Competitive readings highly unlikely 5	Competitive readings highly unlikely 5	Competitive readings highly unlikely 5

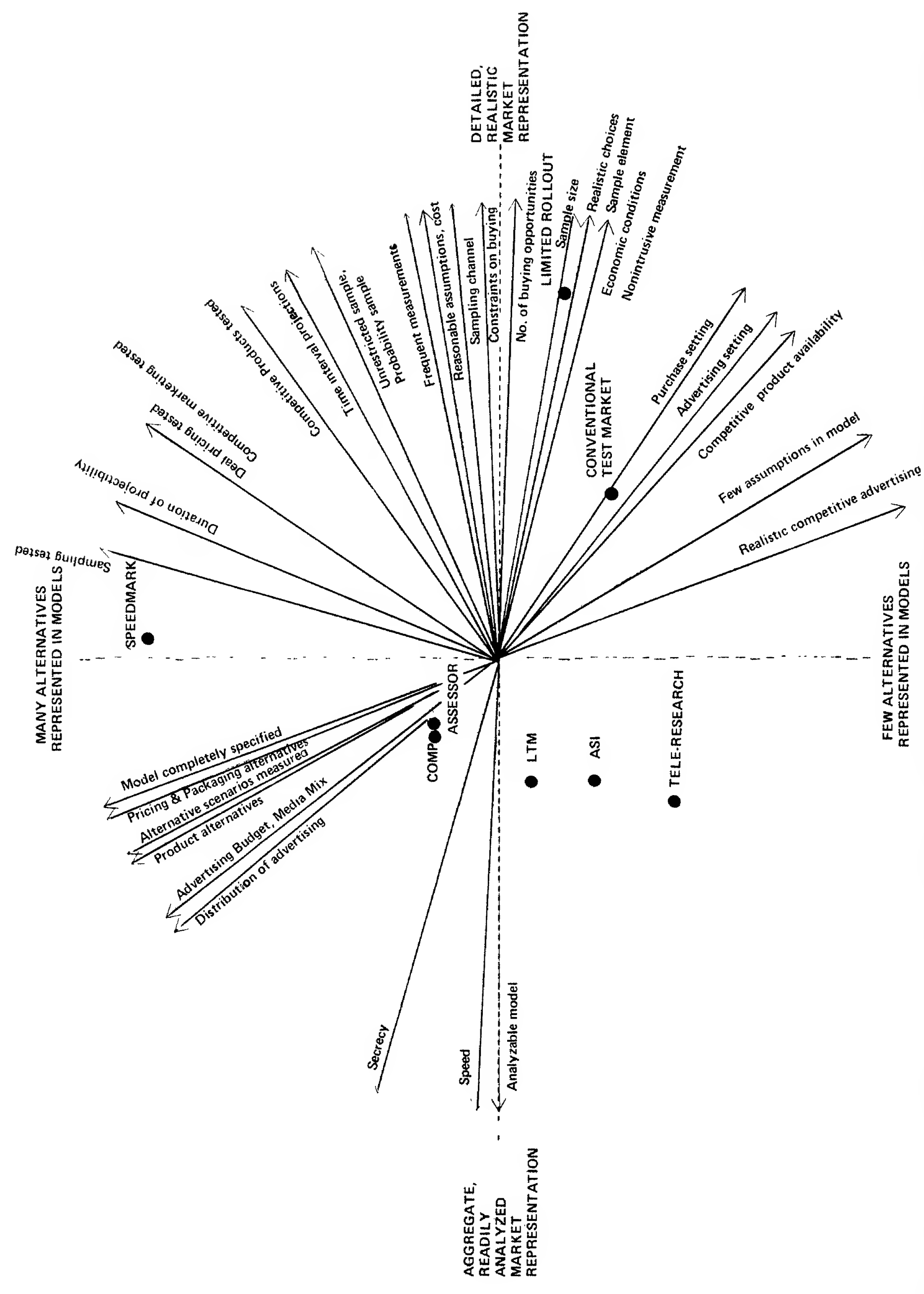


Figure 7–2. Joint-Space Comparative Map of Eight Market-Measurement Procedures and Thirty-Seven Detailed Performance Attributes

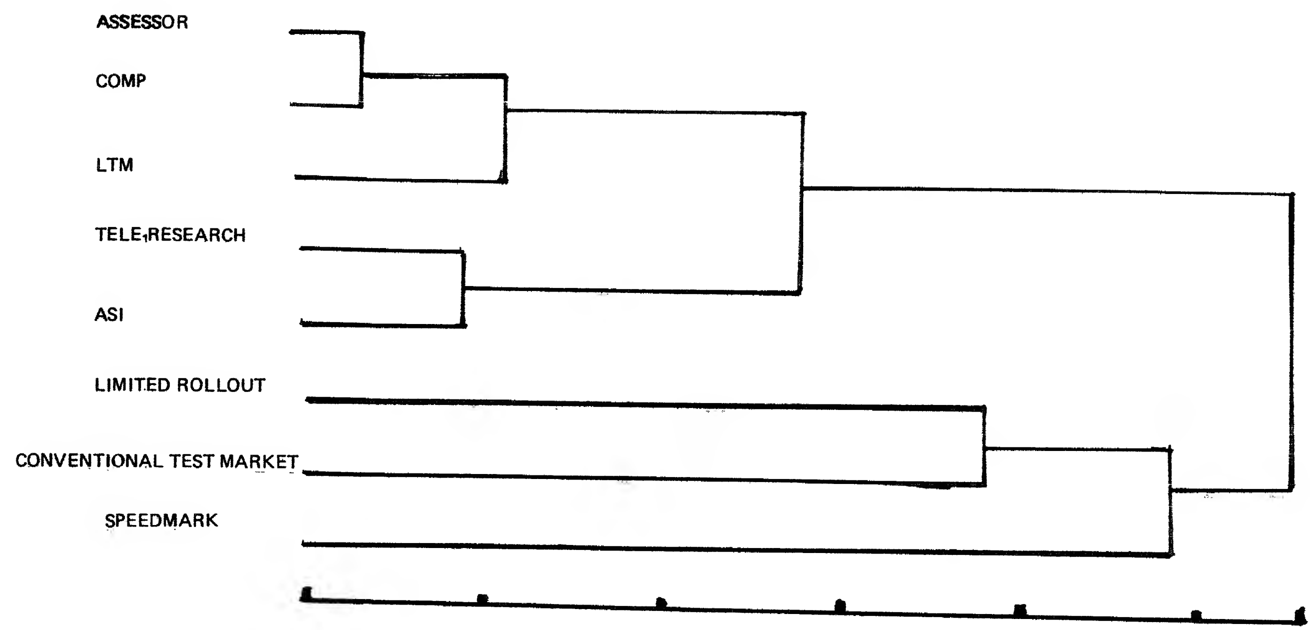


Figure 7–1. Numerical Taxonomy Clustering Diagram of Eight Market-Measurement Procedures Based on Their Thirty-Seven Detailed Attributes

niques in relation to their thirty-seven measurement-attributes vectors. Interpretable graphic representations emerge in the characteristic "family tree" diagram shown in figure 7–1. One such solution is the hierarchical clustering solution presented in figure 7–1.

In this dimensionless portrayal of relative similarities and differences, the closest taxonomic relationships are those which link together near the left side of the tree diagram. A number of interpretive observations follow from an examination of the tree structure, including the following:

1. Conventional test markets and limited rollouts are more closely related to each other than to any simulated test-market procedure and form what may be thought of as the "traditional cluster."
2. One of the simulated test-market procedures, SPEEDMARK® , appears to approach this pairing proximity.
3. The remaining five test-market procedures share sufficient characteristics to result in their residual cluster formation with two distinct subsets—ASSESSOR, COMP, and LTM as one and TELE-RESEARCH and ASI as the other.

In comparing the eight alternative market-measurement procedures on the thirty-seven attributes, a useful visualization of the relationship among the procedures and their attributes is provided using a joint-space map presented in figure 7–2. An examination of the joint-space maps (including the cluster boundaries obtained from the preceding clustering diagrams) reveals that in some respects traditional methods of test marketing and rollouts appear "best." In other important respects, the alternative procedures are revealed as superior. The principal dimensions of these maps can be labeled as "detailed, realistic market simulation" versus "simple, readily analyzed market simulation" and "many versus few alternatives in model."

By dropping perpendiculars from each point to any performance vector of interest, the relative position of each market-measurement approach is revealed on each vector. A market-measurement procedure's projection on an attribute vector reveals that procedure's relative position on each such attribute.

Conclusions

While it has been said that "all comparisons are odious," there are occasions in which they cannot be avoided. To this end, the foregoing analysis can provide an instructive basis for making the desired interpretations or for modifying or repositioning any market-measurement procedure whose characteristics vary from those stated in the underlying data matrix.

Clearly, each measurement procedure presented here may offer relative advantages over others in light of the cost and value of information and the acceptability of both explicit and implicit assumptions. These comparisons are intended to focus on the similarities and differences among the various procedures based on the comprehensive checklist of specifications and judgments presented in table 7-1. They should be supplemented by the potential user's preferences and importance weights for various attributes.

On the one hand, limited rollout and conventional test market have a number of inherent advantages over the test-market alternatives, with rollout being somewhat more advantageous. These advantages can be summarized as providing a detailed, realistic representation of the marketplace, albeit at the price of secrecy, speed, costly expenditures, and at least partial commitment to a product launch.

On the other hand, the test-market alternatives also have advantages over the "real world." They provide secrecy, speed, and lower cost and commitment at the price of realism and detail. Within the set of alternatives, TELE-RESEARCH appears to provide the most realistic competitive context, purchase setting, and advertising setting, with both speed and low cost; however, it provides no method for projecting future sales. Moving from least to greatest model complexity and consideration of alternatives, we find ASI, LTM, COMP, ASSESSOR, and SPEEDMARK®, in that order. SPEEDMARK® is differentiated as a test-market alternative which provides for many alternative scenarios, increased projectability, with enhanced realism and flexibility at the price of greater model complexity and somewhat higher cost.

References

ASSESSOR (Management Decision Sciences, Weston, Mass.)

Silk, A.J., and Urban, G.L. "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology." *Journal of Marketing Research* 15 (May 1978):171-191. (Chapter 8 in this text.)

BASES (Burke Co., Cincinnati)

"BASES—A Management System for Reducing the High Cost and Risk of New Product Development." Sales brochure.

COMP (Elrick & Lavidge, Chicago)

Burger, P.C. "COMP: A New Product Forecasting System" (Northwestern University Working Paper, August 1972). For a more up-to-date version, see chapter 10 in this text.

LITMUS

Blackburn, J.D., and Clancy, K.J. "LITMUS: A New Product Planning Model" (Paper presented at 2d ORSA/TIMS Conference on Market Measurement and Analysis, Austin, Texas, March 1980).

LTM (Yankelovich, Skelly & White, New York)

Goldberg, R. "Basic Description of the Estimating Process for the Laboratory Test Market." Memorandum, May 1973.

MICROMARKET (Tele-Research, Los Angeles)

No published material available.

PURCHASE ACTION (ASI—Audience Studies Inc.—New York)

"Purchase Action." Sales brochure.

SPEEDMARK® (Robinson Associates Inc., Bryn Mawr, Penn.)

"SPEEDMARK Projection System: Technical Methodology." Internal memorandum, June 1975.

General

SSC&B Inc., Research and Information Services Department. "A Review of Outside Research Supplier New Product/Simulation Shopping Testing Services" (July 1976).

8 Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology

Alvin J. Silk and
Glen L. Urban

Introduction

Test marketing is a familiar step in the development of new packaged goods—i.e., branded, low-priced, frequently purchased consumer products. Experimental launchings of new products are intended to expose problems that otherwise would be undetected until full-scale introductions are underway. Although test marketing is commonplace, deciding if and when it should be used in particular cases is a perplexing and controversial management problem. The substantial failure rate historically observed among new packaged goods placed in test markets, plus the high and ever-rising direct cost of such activities, has stimulated firms to seek ways to perform more thorough evaluations of new products *before* embarking on test-marketing programs.

The purpose of this article is to report progress in development of a measurement and model system, called ASSESSOR, designed to estimate the sales potential of new packaged goods before they are test-marketed. The ultimate aim of such a capability is to reduce the incidence of new product failures in test markets and thereby to effect savings in the total cost of new product development.

First some data bearing on test market failure rates and costs are reviewed and current pre-test-marketing evaluation methods are examined briefly. After the particular objectives of ASSESSOR are set forth, the measurement methodology, model structure, and estimation procedures

---

Reprinted from A. Silk and G. Urban, "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research* 15 (May 1978): 171–191, published by the American Marketing Association.

The authors are indebted to Calvin Hodock for important discussions of this subject and essential marketing research support. Special acknowledgement is due Robert Klein who contributed many valuable suggestions and much expertise in data collection and analyses to this project. Helpful comments on a draft were given by John Hauser and two *JMR* referees.

used are described. The first application of the system then is discussed in some detail. Finally, some results obtained from subsequent studies are reviewed briefly and the limitations of the approach are considered.

Problem Description

Test Marketing

Manufacturers of packaged goods have come to rely on a fairly common set of measurement methods for assessing consumer response to a new product. The typical approach includes (1) concept and positioning tests, (2) product-usage tests, and (3) test marketing [34; 47 chapter 4]. The last step constitutes the final integration and evaluation of the product formulation and the various elements of the marketing plan designed to implement the desired positioning strategy.

The design and scale of test market operations for new products depend on the specification of purpose in terms of estimation and experimentation [42]. The objective of test marketing is sometimes primarily to obtain an estimate of the market share and/or sales volume that would be realized if the new product were launched nationally. In other cases the aim may be to evaluate alternative marketing mix strategies, and hence the test marketing program involves a true experiment. A recent survey of the test marketing practices of "28 major consumer grocery and drug product companies" in the U.S. found that the "norm" was to run a test market in three areas for 10 or 11 months [33]. Over a three-year period, these firms had averaged three test marketing programs each per year. The costs of such efforts are considerable and have been mounting. In 1967 the "going rate" for a year-long test in several markets was reported to be $500,000 [61, p. 45]. Today the comparable figure appears to be nearly a million dollars and the authors are familiar with several three-city test marketing programs that involved outlays of $1,500,000.

Even more than the costs, what has motivated closer scrutiny of test marketing practices is recognition of the distressingly high probability that such an undertaking will lead to the detection of a new product failure rather than a success. A review of the limited data available suggests that either outcome is equally likely. In 1961 and again in 1971, the A.C. Nielsen Company reported the "success ratio" of new brands (health and beauty aids, household and grocery products) that had been test marketed through their facilities [50]. The 1961 study included 103 new brands and the 1971 covered 204 items. "Success" was defined by the "manufacturer's judgment of each brand's performance in test"—namely whether or not the brand was launched nationally. Brands withdrawn from test markets or not introduced nationally were considered "failures." By these criteria, only

about half of the new brands test marketed in these two periods were successes (54.4% in 1961, 46.6% in 1971). Similarly, the aforementioned survey of the test marketing practices of 28 major consumer grocery and drug product companies found that in 46% of the 54 specific test market experiences covered by the study, test market sales "fell short of management expectations" [33]. In contrast, Buzzell and Nourse [16, p.100] observed in their study of the food industry that only 32% of 84 "distinctly new food products" developed in the 1954–1964 period were discontinued after test marketing. This somewhat lower failure rate is probably related to the special character of the sample of products studied—i.e., all were "substantially different in form, ingredients, or processing methods from other products previously marketed by a *given* company" [16, p.96]. At the individual firm level, 10-year test market success rates of 46% and 60% have been reported for General Foods in the U.S. [5, p. 50] and Cadbury in the U.K., respectively [17, p. 98]. Thus, failure rates ranging from 40 to 60% roughly bracket the publicly reported record of test market experience in the packaged goods field.

Besides being an expensive means of detecting new product failures, test marketing involves other problems. First, the test market performance of a new product can be monitored by competitors and provides them with information and time needed to plan a response. Second, the external validity or "projectability" of test market results to subsequent national performance has long been a subject of debate and controversy [1]. For example, A.C. Nielsen Company compared the first year national market share position of 50 new brands with their first year test market performance and concluded that "the odds are about 50–50 that the national performance will match test results within ±10%" [51, p. 4]. This kind of straightforward comparison assumes that test market conditions with respect to such factors as promotional and distribution support and competitive activity were representative of circumstances in the national market. Such an assumption is rarely, if ever, tenable and experience indicates that the predictive accuracy of test-market-based forecasts can be improved markedly by adjusting for discrepancies between test and national conditions with the help of a model that accounts for the dynamics of the new product response process. Competitors have been known to take deliberate retaliatory actions to disrupt another firm's test markets which make it extremely difficult to untangle the results even by complex, model-based analyses [66].

Pre-Test-Market Evaluation

Packaged goods manufacturers have sought in a variety of ways [17] to cope with the high incidence of new product failures in test markets, diffi-

culties in projecting test market results, and the heavy cost of such activities. The most logical place for improvements is the early stages of the new product development process. More effective search and screening procedures can increase the productivity of development and test efforts and diminish the likelihood that a failure will not be detected until the test marketing stage. In recent years new measurement methods and models have been developed and used to facilitate concept generation, refinement, and evaluation [31, 53, 60, 62, 72]. Application of these techniques is intended to lead to better concepts and products, but does not ordinarily obviate test marketing.

Attention also has been directed to making more careful forecasts of expected test market results before the launching of such operations. The sales and market share observed over time for new, frequently purchased consumer products tend to follow a consistent general pattern that can be understood in terms of the level of cumulative trial the new brand achieves and the rate of repeat purchasing it is able to sustain [52]. As the diffusion process proceeds, trial and repeat purchases move toward steady-state levels giving rise to an equilibrium market share and sales rate. Several models have been developed which use early test market or introductory data to forecast equilibrium share and volume [41, Chapter 17]. Certain of these models have been used to arrive at pre-test-market predictions of equilibrium share employing inputs derived from concept and usage tests for the new product plus data for analogous products and/or judgment [68, 69]. Claycamp and Liddy [22] carried this idea a step further and built a regression model to predict trial and repeat purchase levels before the launch of a test market from a set of controllable and uncontrollable variables measured by a mixture of judgmental ratings and consumer test results. The model was estimated and tested with data obtained from 58 new product introductions that covered 32 different types of packaged goods. Eskin and Malec [28] report progress in developing a model which extends the Claycamp and Liddy work in important ways. Some firms have developed similar models using historical data on new product introductions for more narrowly defined product categories [6; 47, p. 94–100]. Though the evidence reported bearing on the forecasting ability of this approach is encouraging [6, 22, 69], the use of such cross-sectional models is always surrounded by uncertainty about the universe of new products and market conditions over which the parameter estimates can be expected to remain stable [26, 64].

Many packaged goods manufacturers have turned to lower cost alternatives to the traditional multi-area test market as a means of reducing expenditures on new product research [e.g., 17]. Several varieties of scaled-down or "controlled" test markets have been used [1]. These operations typically involve fewer and/or smaller areas, but allow more control over

some marketing mix variables than do regular test markets. However, the costs remain substantial (expenditures of $100,000 are common) and the projectability of results to the total market is controversial [47, p. 48]. A related but essentially different approach operative in Western Europe, the "mini test market" [29, 57], is discussed briefly in the section on design considerations.

Another pre-test-market method for evaluating new packaged goods is the "laboratory" or "simulated" test market. The basic design concept is to simulate the awareness-trial-repeat purchase process by controlled laboratory and product usage tests. Measurements obtained at several points in this process are used to predict steady-state market share for the new brand and to provide diagnostic information. These ideas form the basis of the work reported here. Brief mention of previous applications of this type of combined laboratory-use test design in commercial marketing research can be found in the literature [47, p. 44, 59; 49, p. 77–9, 183–5; 64], and several firms are known to offer such services [64]. However, the only detailed account of comparable work known to the authors is in an unpublished paper by Burger [15] who describes the COMP system developed in conjunction with Elrick and Lavidge, Inc. The specific measurements, models, and estimation procedures used in the present study are very different from Burger's.

Objectives and Structure of ASSESSOR

ASSESSOR is a set of measurement procedures and models designed to aid management in evaluating new packaged goods before test marketing when a positioning strategy has been developed and executed to the point where the product, packaging, and advertising copy are available and an introductory marketing plan (price, promotion, and advertising) has been formulated. Given these inputs, the system is intended to:

1. Predict the new brand's equilibrium or long-run market share.
2. Estimate the sources of the new brand's share—"cannibalization" of the firm's existing brand(s) and "draw" from competitor's brands.
3. Produce actionable diagnostic information for product improvement and the development of advertising copy and other creative materials.
4. Permit low cost screening of selected elements of alternative marketing plans (advertising copy, price, and package design).

Figure 8-1 shows the overall structure of the system developed to meet these requirements. The critical task of predicting the brand's market share is approached through two models—one relates preference to purchase

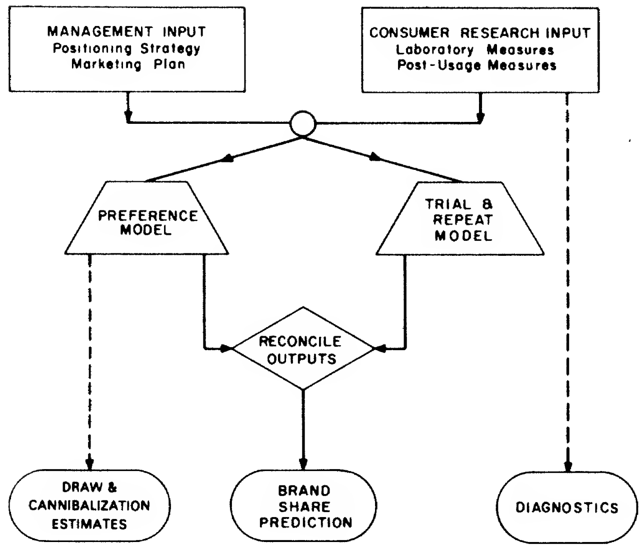


Figure 8–1. Structure of the ASSESSOR System

probability and the other is a straightforward flow representation of the trial-repeat process. The two models are similar in structure, but are calibrated in different ways. Convergent results should strengthen confidence in the prediction whereas divergent outcomes signal the need for further analyses to identify sources of discrepancies and to provide bases for reconciliation. The measurement inputs required for both models are obtained from a research design involving laboratory and usage tests. The key outputs are a market share prediction plus diagnostic information which can be used to make a decision as to the brand's future. Several outcomes are possible. A poor showing may lead to either termination or further developmental efforts. If the performance is satisfactory, plans for test marketing can proceed. Very favorable results could lead to an immediate launching of the brand, particularly if the capital investment risked in the introduction is small and/or if the threat of competitive entry is imminent.

Research Design and Measurement

An Overview of the Design

The measurement inputs required to develop the desired diagnostic information and predictions for ASSESSOR are obtained from a research design structured to parallell the basic stages of the process of consumer response

Table 8–1
Research Design and Measurement

Design	*Procedure*	*Measurement*
Q_1	Respondent screening and recruitment (personal interview)	Criteria for target-group identification (for example, product-class use)
Q_2	Premeasurement for established brands (self-administered questionnaire)	Composition of "relevant set" of established brands, attribute weights and ratings, and preferences
X_1	Exposure to advertising for established brands *and* new brand	
[Q_3]	Measurement of reactions to the advertising materials (self-administered questionnaire)	Optional, for example, likability and believability of advertising materials
X_2	Simulated shopping trip and exposure to display of new and established brands	
Q_4	Purchase opportunity (choice recorded by research personnel)	Brand(s) purchased
X_3	Home use/consumption of new brand	
Q_5	Postuse measurement (telephone interview)	New-brand use rate, satisfaction ratings, and repeat purchase propensity: attribute ratings and preferences for "relevant set" of established brands plus new brand

Q = measurement
X = advertising or product exposure

to a new product. Table 8–1 outlines the essential features of the design and identifies the main types of data collected at each step. To simulate the awareness-trial stages of the response process, a laboratory-based experimental procedure is used wherein a sample of consumers are exposed to advertising for the new product and a small set of the principal competing products already established in the market. Next, the consumers enter a simulated shopping facility where they have the opportunity to purchase quantities of the new and/or established products. The ability of the new product to attract repeat purchases is assessed by one or more waves of followup interviews with the same respondents conducted after enough time has passed for them to have used or consumed a significant quantity of the new product at home.

Procedures

The laboratory phase of the research is executed in a facility in the immediate vicinity of a shopping center. "Intercept" interviews (Q_1) are con-

ducted with shoppers to screen and recruit a sample of consumers having attributes that characterize the target market for the new product. The schedule of this work is staggered over time to reduce the opportunity for obvious kinds of self-selection biases to affect the respondents drawn into the study. Additional control over sample composition can be exercised by carrying out the field work at several different locations chosen to attain the heterogeneity and quotas desired in the final sample. Considerable flexibility is possible here because elaborate facilities and arrangements are not required. Studies done to date typically have involved samples of approximately 300 persons.

Upon arriving at the laboratory facility, respondents are asked to complete a self-administered questionnaire that constitutes the before measurement (Q_2). Individually or in pairs, respondents then proceed to a separate area where they are shown a set of advertising materials (X_1) for the new brand plus the leading established brands. Ordinarily, respondents are exposed to five or six commercials, one per brand, and the presentation order is rotated for different groups to avoid any systematic position effects. Measurement of reactions to the advertising materials (Q_3) is done next if such information is desired for diagnostic purposes. Dropping this optional feature of the design eliminates a potential source of unwanted reactive effects on respondents' subsequent behavior.

The final stage of the laboratory experiment takes place in a simulated retail store where participants have the opportunity to make a purchase. When first approached, they are told that they will be given a fixed amount of compensation for their time—typically about \$2.00 but always more than the sum needed to make a purchase. In the lab they are informed that they may use the money to purchase any brand or combination of brands in the product category they choose and that they can keep any unexpended cash. They then move to an area where quantities of the full set of competing brands including the new one are displayed and available for inspection (X_2). Each brand is priced at a level equal to the average price at which it is being sold regularly in mass retail outlets in the local market area. The brand (or brands) selected by each participant is (are) recorded by one of the research personnel (Q_4) at the checkout counter. Although respondents are free to forego buying anything and to retain the full \$2.00, most do make a purchase. For example, the proportions of participants making a purchase observed in two separate studies of deodorants and antacids were 74% and 64%, respectively. Those who do not purchase the new brand are given a quantity of it free after all buying transactions have been completed. Note that this procedure parallels the common practice of effecting trial usage by the distribution of free samples. A record is maintained for each respondent as to whether he or she "purchased" or was given the new brand to allow assessment of whether responses on the post-usage survey are affected differentially by trial purchase versus free sampling.

The post-usage survey (Q_5) is administered by telephone after enough time has passed for usage experience to have developed. The specific length of the pre-post measurement interval is determined by the estimated average usage rate for the new product. Respondents are offered an opportunity to make a repurchase of the new brand (to be delivered by mail) and respond to essentially the same set of perception and preference measurements that was used in the before or pre-measurement step (Q_2) except that they now rate the new brand as well as established ones. Familiarity with the questionnaire gained from previous exposure makes it feasible to re-administer the instruments in telephone interviews.

Some shrinkage in sample size inevitably occurs between the laboratory session and the post-usage survey. Two important varieties of attrition occur. First, some proportion of persons who participated in the laboratory session will be excluded from the telephone survey as a result of having moved, being away from home, refusing to be interviewed, etc. A second source of sample attrition is respondents who report in the post-usage survey that they have not used the supply of the new product they previously had purchased in the lab store or had been given. In the deodorant study, 16.7 % of the original laboratory sample could not be reinterviewed, and another 16.7% had not used the product. The general policy has been to continue reinterview efforts until a sample of users of the new product is obtained which includes at least two-thirds of the original set of respondents. Subjects not responding to the post-usage survey can be compared with those who do with respect to information about such factors as last purchase brand share and usage rate obtained from the before measurement (Q_2) in order to detect the presence of systematic biases in the post-usage sample due to experimental attrition.

Measurement Instruments

Table 8–1 identifies the key measures obtained at various points in the design. Certain nonstandard features of the methods require additional discussion. Allaire [3] has shown that measurement of perception and preference structures can be distorted by including unfamiliar stimuli in the set of alternatives judged. Following his methodological recommendation, the authors ask each respondent to provide perception and preference ratings only for those brands that compose his or her "relevant set" of alternatives—i.e., that subset of available brands which are familiar to the respondent regardless of whether they are judged favorably or unfavorably as choice alternatives.[1] Respondents' idiosyncratic relevant sets are revealed by a series of unaided recall questions which identify brands previously purchased or used plus any others considered to be satisfactory or unsatisfactory alternatives.

The size of a typical respondent's relevant set is small in relation to the total number of brands available in the market. Urban's [67] data for seven different categories of packaged goods show that the median relevant set size generally observed is about three brands. Campbell [18] and Rao [59] reported evoked set sizes of approximately the same magnitude for some additional product classes. The smallness of evoked or relevant set sizes is consistent with evidence available as to the number of different brands of packaged goods actually purchased by households. Massy et al. [44, pp. 22–4] reported some relevant statistics for a subsample of U.S. households in the J. Walter Thompson panel. During a one-year period, the mean number of different brands purchased per household was 3.3 for regular coffee, 2.6 for tea, and 3.0 for beer. The ranges observed in this quantity for these three product categories were 1 to 12, 1 to 8, and 1 to 11, respectively. Wierenga [70, chapter 6] investigated some related phenomena using purchase diary data from a panel of 2,000 Dutch households. He found that although a total of 29 different brands accounted for 85% of the total volume of margarine purchased, the mean number of brands purchased per household over a two-year period was only 4.26. The comparable figures for beer and an unidentified food product were eight and 14 brands available, respectively, with 2.57 and 2.88 being the average number of brands purchased per household in these two product categories. Table 8–2 shows the distribution of relevant set sizes for deodorants observed among a sample of 299 respondents. Here again, the median relevant size is three brands.

After a respondent's relevant set of brands is identified, attribute importance ratings are obtained. Beliefs/perceptions about the extent to which each brand in a respondent's relevant set offers these attributes also are elicited by means of bipolar satisfaction scales. These two types of data are important to components of the diagnostic information provided by the system.

A constant sum, paired comparison procedure is used to assess brand preferences. Several variants of the constant sum approach have been used in marketing research studies and some evidence bearing on the reliability and validity of such measures has been reported. Axelrod [7] employed a constant sum technique as a rating scale device by asking respondents to allocate "11 cards" among a predetermined set of brands so as to indicate the likelihood of their buying each brand. An individual's preference score for a particular brand was simply the number of cards allocated to it. In a complex multistage study, several different awareness and preference measures were compared with respect to their "sensitivity" (ability to detect an effect of advertising exposure in a before-after with control group design), "stability" (aggregate agreement between equivalent samples), and "predictive power" (ability to predict purchases at t_2 from measures ob-

Table 8–2
Distribution of Relevant Set Sizes for Deodorants

Relevant Set Size	*Percentage of Sample (N = 299)*
1 or 2	31.8
3	31.8
4	23.1
5	7.0
6	4.0
7	2.3
	100.0

tained at t_1). On the basis of the results obtained, Axelrod recommended use of the constant sum scale to elicit attitude ratings for brands mentioned by consumers in response to an unaided brand awareness question.

Haley [32] reported the results from another comparative study of several attitudinal measures which included a combined paired comparison, constant sum procedure. For all possible pairs of brands, respondents were instructed to divide "10 points" between any two brands so as to reflect their preferences. An individual's preference score was obtained for each brand by summing the points assigned to that brand over all the relevant pairwise comparisons. In relation to the other measures investigated, Haley reported that this method proved superior in its ability to discriminate among brands. Also, it yielded scores whose distribution appeared to be approximately normal.

The findings reported by Axelrod and Haley suggested use of the constant sum technique as a desirable procedure for eliciting preference judgments from consumers. However, in both of these studies and in other marketing-research applications, the methods used to estimate scale values for brands from constant sum input data have been of an *ad hoc* variety. In psychophysical measurement where it was first used [65, pp. 105–7], constant sum comparative judgments are the basis of an explicit scaling model for which formal estimation methods have been developed. Under the assumption that the subjects can provide *ratio* judgments of paired comparisons between stimuli, Torgerson [65, pp. 108–12] devised a least-squares method for estimating ratio scale values. The authors used this form of constant sum, paired comparison scaling to measure a respondent's preferences for his or her relevant set of brands.

The measures of attribute importance weights, brand belief or attribute

ratings, and preferences obtained in the before measurement (Q_2) are repeated again in the post-usage survey (Q_5) but with the new brand added to each respondent's "relevant set" of alternatives. Finally, respondents are given an opportunity to make a mail order repurchase of the new product.

Design Considerations

Selection of the design outlined was influenced by certain operational cost and timing objectives. In particular, the new product management group which initiated this work was seeking a method of producing an evaluation of a new packaged good within a three-month period and at a cost of less than 5% of the typical expenditure required for a test market (i.e., $25,000 to 50,000). The time and expense required to implement the data collection procedures described here are well within the limits of these design desiderata.

An additional appealing feature of the design is flexibility. It can be expanded for a relatively modest amount of incremental cost to permit evaluations of alternative executions of certain elements of the new product's introductory marketing program—e.g., responses to different commercials can be compared by adding treatment groups to the design, each of which is exposed to a separate commercial.

Mail and home delivery panels are possible alternatives to the approach described. The difficulty of efficiently reaching respondents from the relevant target group and the problem of nonresponse diminish the attractiveness of mail panels. The authors are aware of no published accounts of the use of mail panels for testing new packaged goods. Evidence of the successful use of a home delivery panel in new product testing has been reported by Pymont and his coworkers [21, 29, 57, 58]. Initially developed in the United Kingdom as a "mini test market" facility and subsequently adopted in several other Western European countries, this carefully conceived measurement system involves a continuous panel of households who make purchases from a special door-to-door retail grocery service. Promotional communications and new product introductions are effected by means of controlled print vehicles sent to members of the panel. Charlton et al. [20] analyzed the purchase behavior observed in this environment using Ehrenberg's [25] NBD repeat buying model. They concluded that the brand choice patterns of the mini test panel for established products "are generally like those in real life" in the sense of being consistent with models known to describe purchase behavior under natural conditions. This methodology has been used extensively in Western Europe to evaluate new packaged goods, the obvious attraction being that it offers an efficient

means of estimating repeat purchasing for new brands. The latter quantity generally is acknowledged to be the prime determinant of a new brand's success or failure, but the one which is least amenable to rapid and accurate measurement. A high degree of predictive accuracy is claimed for this system and supported by case histories of several applications. Steady-state shares predicted for new brands by the Parfitt-Collins [52] model using estimates of the trial, repeat, and buying rate parameters derived from the mini test panel have been found to be in very close agreement with the comparable share figures observed in concurrent or subsequent normal test markets and/or national introductions [57, 58].

The home delivery panel/mini test market clearly is an appealing alternative to the approach pursued here. It remains an open question whether the former methodology ordinarily would allow the timing and cost criteria established for the present work to be met. The authors have found no documented accounts or other reports of experiences in the U.S. with a mini test market facility. In the United Kingdom, expenditures for new product penetration studies in the mini test market system are said to be "less than 5% of the cost of conventional test marketing" and, on the average, about 16 weeks of testing is required [29]. The home delivery arrangement does not permit the television commercial and product display exposure that can be effected in a laboratory facility and hence trial usage can be expected to accumulate more rapidly under the latter approach. However, if extensive usage experience is required for consumers to learn about the new product or if its frequency of purchase differs from that of established brands [8, 27], then one or two waves of post-usage interviews conducted soon after the laboratory session will not provide a reliable basis for estimating its repeat buying rate, and the home delivery panel becomes a preferred and necessary alternative.

Cost and timing considerations aside, the larger design issue is the quality of the measurements. Reference was made heretofore to various steps taken to minimize and/or identify certain threats to validity [19]. One other potential source of confounding effects that merits attention is the use of repeated measures with the same respondents. The available evidence suggests that this is not a troublesome feature of the present design. In a special experimental study undertaken to investigate this issue, the measurement of response to the advertising materials (O_3) was found to have no apparent reactive effect on respondent brand choice behavior observed (Q_4) in the simulated shopping trip. Ginter [30] and Winter [73] also investigated this general issue in their laboratory study which involved four consecutive weekly sets of measurements taken before and after exposure to advertising stimuli. They found some indications that the repeated measurements were reactive, but report that these effects were not sufficiently strong or systematic to be problematical. [30, p.33; 73, p. 32].

Model Structure

As shown in figure 8–1, two different models are used to generate separate predictions of market share for a new brand. The first relates strength of post-trial preference for the new brand to the probability of purchasing it. The second is a more direct representation of the trial-repeat purchase process. In this section the details of each model are set forth and their structural correspondence is examined. A discussion of how the output of the system is used for strategic management issues follows.

Preference Model

The fundamental problem addressed here is that of predicting market share, an aggregate measure of purchase behavior. The available empirical evidence leads one to favor selection of preference over other attitudinal or behavioral disposition constructs as a simple predictor of brand choice. As was mentioned, Axelrod [7] found the predictive power of preference ratings (obtained by a constant sum procedure) superior to that of a variety of other interview/questionnaire-based evaluative measures for established brands of packaged goods.

In the first of a series of important studies, Pessemier et al. [55] demonstrated that their interval-scaled "dollarmetric" measure of brand preferences [54] obtained in a laboratory setting could be used to develop fairly accurate predictions of the relative frequency of individual consumer's subsequent purchases of established brands under natural conditions over a seven-month period. More recently, Ginter [30] conducted an experimental study of response to a new brand that involved a sequence of four weekly laboratory sessions wherein housewives were exposed to commercials for new brands in two different packaged goods categories and were given the opportunity to purchase them in a simulated shopping trip. Among other things, he found that preference (measured by the same method as that used previously by Pessemier et al.) was a better predictor of purchase of the new brands than a multiattribute model. The several unresolved issues [71] that surround the latter class of models further discourage their use for the present purposes.

On the basis of subsequent work, Bass et al. [11] argue that although preference measures do exhibit significant predictive power, a high degree of accuracy cannot be realized because of measurement error, omitted variables, random exogenous events, etc., and perhaps consumers' "desire for variety". This leads them to the view that "since choice behavior is not constant even when attitudes are unchanging, attitude-based predictions of choice must be probabilistic" [11, p. 541]. (Bass has gone further in developing this position. See Ref. 9).

A similar orientation has been adopted here: the authors first estimate individual consumers' probabilities of purchasing the new brand from their expressed brand preferences after a period of initial usage of it and then aggregate these probabilities across individuals to obtain an estimate of expected aggregate or total market share.

Luce's probabilistic theory of choice [43] provides a valuable foundation for formulating a model to link brand preferences to purchase probabilities. Luce has shown that a simple but powerful axiom about choice probabilities implies the existence of a ratio scale for the alternatives. More specifically, the Luce model, written in terms of brand choice probabilities and preferences asserts that:

$$P_i(j) = \frac{V_i(j)}{\sum_{k=1}^{m_i} V_i(k)}, \qquad V_i(k) > 0, \tag{8.1}$$

where: $P_i(j)$ = probability that consumer i chooses brand j,

$V_i(j)$ = consumer i's ratio scaled preference for brand j,

$k = 1, \ldots, j, \ldots, m_i$,

m_i = number of brands in a respondent's relevant set of alternatives.

In the present context, the authors postulate that the observed measures of preference, obtained by the constant sum, paired comparison procedure referred to previously, are related to brand choice probabilities by:

$$P_i(j) = \frac{[\hat{V}_i(j)]^\beta}{\sum_{k=1}^{m_i} [\hat{V}_i(k)]^\beta}, \qquad V_i(k) > 0, \tag{8.2}$$

where: $\hat{V}_i(j)$ = estimated preference of consumer i for brand j,

β = parameter to be estimated.

This form of preference model has been used in consumer research by Pessemier et al. [55]. They found that straightforward application of their interval scale preference measure to equation 8.1 resulted in overprediction of the relative frequency of purchase of less preferred brands. Better fits were realized with equation 8.2 where a heuristic method was used to obtain an estimate of β that was product class specific, but which applied across all brands and all consumers. Pessemier et al. [55] discussed the application of the exponent β to the preference scores as a means of accounting for

noise and discrepancies between laboratory and market conditions. Similarly here, the ratio scaling of preferences to which the constant sum, paired comparison procedure aspires may not be attained and the exact properties of the preference scale rendered by use of the method cannot be directly ascertained. Pessemier and Wilkie [56] pointed out that the transformation implied in equation 8.2 that equates it to equation 8.1 is similar to Steven's Power Law [63] used in psychophysical research to relate subjective magnitude to physical magnitude.

Formulation 8.2 also may be related to McFadden's [46] "random utility model" which he derived as a theory of population choice behavior, building upon Luce's individual choice model. McFadden assumes the utility (μ_{ij}) each member (i) of a utility-maximizing population of consumers has for a choice alternative (j) consists of a measurable (c_{ij}) component and a stochastic element (ϵ_{ij}), i.e.:

$$\mu_i(j) = c_i(j) + \epsilon_i(j). \tag{8.3}$$

The nonstochastic component is taken to be a function of a vector of attributes describing the alternatives faced by the individual. Assuming the $\epsilon_i(j)$ are independent Weibull distributed, McFadden shows that Luce's model of individual behavior leads to an econometric specification of the choice probabilities as a multinomial logit model similar to equation 8.2, i.e.,

$$P_i(j) = \frac{\exp\,[c_i(j)]}{\sum_k \exp\,[c_i(k)]}.$$

Empirical experience also has led to the use of equation 8.2 in this work. The authors estimate β using the preference scale values for the *established brands* derived from data obtained in the pre-exposure questionnaire (Q_2 in table 8-1) and information about the last brand which respondents report having purchased. Statistical methods for estimating β are discussed in the next section. Assuming β to be a stable parameter whose value will remain unchanged after introduction of the new brand, and given measures of consumers' preferences for the new brand plus the established brands obtained *after* a period of trial usage of the new brand, it follows from equation 8.2 that one can predict each individual's probability of purchasing the new brand using:

$$L_i(t) = \frac{[A_i(t)]^\beta}{[A_i(t)]^\beta + \sum_{k=1}^{mi} [A_i(k)]^\beta}, \tag{8.4}$$

where $L_i(t)$ = probability that consumer i chooses the brand t after having tried the new brand,

t = index for the new brand,

k = index for established brands,

$A_i(t)$ = estimated preference of consumer i for the new brand t after having tried the new brand,

$A_i(k)$ = estimated preference of consumer i for established brand k after having tried the new brand.

Now the predicted probabilities are conditional upon the brand being an element of each consumers' relevant set. To calculate an expected market share for the new brand one must take into account that the new brand will not necessarily become an element of the relevant set of brands for all consumers when it does become available in the market. Therefore,

$$M(t) = E(t)\frac{\sum_{i=1}^{N} L_i(t)}{N}, \tag{8.5}$$

where: $M(t)$ = expected market share for the new brand t,

$E(t)$ = proportion of consumers who include brand t in their relevant set of alternatives,

$L_i(t)$ = predicted probability of purchase brand t by consumer i, i = $l, \ldots, N$.

To use equation 8.5 to forecast the new brand's market share, one first must predict the proportion of consumers, $E(t)$, who will consider the new brand as a relevant alternative. A procedure for estimating this quantity is discussed in the next section.

Where there is substantial variation in consumption among individual consumers, the $L_i(t)$ in equation 8.5 must be weighted by a usage rate index.

The task of predicting how the new brand will affect the shares of existing brands requires that one obtain their expected market share when equilibrium is reestablished after the launching of the new brand. To do so, one must again recognize that under the new steady-state conditions the market will consist of two subpopulations, distinguishable by the presence or absence of the new brand in their relevant sets. The sizes of these two groups in relation to the total target market will be $E(t)$ and $1-E(t)$, respectively. The addition of the new brand to a respondent's relevant sets is effected experimentally by the procedures noted previously and so the impact of its inclusion will be manifested in the preferences for the established brands expressed by respondents in the post-usage survey (Q) *after* exposure to the new brands—i.e., in the quantities $A_i(k)$. In contrast, it seems

reasonable to suppose that consumers whose relevant set does not include the new brand will continue to purchase established brands after the new brand is available in the same manner as they did before its entry—i.e., according to the established brand preferences held *before* exposure to the new brand, $\hat{V}_i(k)$, as defined in equation 8.2. One also assumes that (1) the probability of the new brand being included in a consumer's relevant set is independent of relevant set size and composition or the structure of preferences for established brands, and (2) inclusion of the new brand in a consumer's relevant set does not affect the number or identity of established brands it contains. Using these ideas one derives expected market shares for established brands in the following manner. As in equation 8.4, if the new brand is present in a consumer's set, the purchase probability for any established brand j will be given by:

$$L_i(j) = \frac{[A_i(j)]^\beta}{[A_i(t)]^\beta + \sum_{k=1}^{m_i} [A_i(k)]^\beta}, \tag{8.6}$$

$$k = 1, \ldots, j, \ldots, m_i,$$

and its share in the submarket of consumers whose relevant set includes the new brand is:

$$M(j)' = \frac{\sum_{r} {}_{r(j)} L_i(j)}{N}, \tag{8.7}$$

where the summation $\Sigma_{r(j)} L_i(j)$ is over the $r(j)$ individuals who include the established brand j in their relevant sets.

For consumers who do not include the new brand in their relevant sets, the probability of purchasing any established brand j can be obtained from equation 8.2 and within the subpopulation of all such consumers its market share will be:

$$M(j)'' = \frac{\sum_{r(j)} P_i(j)}{N}, \tag{8.8}$$

To obtain the established brand's expected market share in a total market, one weights the "unadjusted" shares (equations 8.7 and 8.8) by the relative sizes of the two subpopulations, or:

$$M(j) = E(t)M'(j) + (1 - E(t))M''(j), \tag{8.9}$$

where: $M(j)$ = expected market share for the established brand j after introduction of the new brand t.

Note that because it follows from equations 8.4 and 8.6 that:

$$\sum_{k=1}^{m_i} L_i(k) + L_i(t) = 1, \qquad i = 1, \ldots, N, \tag{8.10}$$

and similarly because from equation 8.2:

$$\sum_{k=1}^{m_i} P_i(k) = 1, \qquad i = 1, \ldots, N, \tag{8.11}$$

then the expected market shares given by equations 8.5 and 8.8 will be logically consistent:

$$\sum_{k=1}^{m^*} M(k) + M(t) = 1, \tag{8.12}$$

where: m^* = total number of existing or established brands.

Comparing the expected market share given by equation 8.9 for any established brand with its prior share permits one to estimate the impact of the new brand on the existing structure of market shares for established brands.

The key assumptions underlying the model serve to emphasize the conditions under which it can be expected to apply. First, crucial to the Luce-McFadden choice models is the notion of "independence of irrelevant alternatives." Formally, the requirement is that the "ratio of the probability of choosing one alternative to the probability of choosing the other should not depend upon the total set of alternatives available" [43, p. 9]. As discussed elsewhere [23, p. 150–1; 46, p.113], this assumption will not hold when the set of alternatives is sufficiently heterogeneous that choices are made in a hierarchical manner, as when a consumer first selects among several product types and then chooses a brand within a particular subcategory. The practice followed here of identifying idiosyncratic relevant sets of alternatives appears to offer some protection against mixing together alternatives that differ markedly in their perceived substitutability. Though it may also be possible to model the structure of a hierarchical choice process separately, attention ultimately must be focused on relatively homogeneous sets

of alternatives. Some evidence bearing on the independence-of-irrelevant-alternatives assumption will be discussed in connection with estimation of the preference model.

A second important assumption is the treatment of brand choice as a heterogeneous, stationary, zero-order Bernoulli process [45, chapter 3]. A survey of the issues and pertinent evidence is given by Bass [10] who emphasizes that stochastic choice models built on these premises are consistent with stable market shares accompanied by substantial brand switching, conditions which are frequently observed together in packaged goods markets. Bass et al. [12] discuss the relationship between heterogeneous, zero-order brand switching models and penetration models like those of Ehrenberg and his coworkers [25]. In addition, they show that, under certain assumptions about how brand preferences are distributed in the population, the Luce choice model leads to a flexible and tractable distribution of purchase probabilities.

Rather than model and measure the dynamics of the adoption process directly, the authors seek to compare equilibrium or steady-state market shares before and after introduction of a new brand, while allowing for heterogeneity in the population of consumers. For the approximation of stationarity to be plausible, market shares for established brands should be constant before the new brand's launch and preferences must have stabilized when the post-usage measures are taken. The latter condition can be checked by repeating the post-usage survey after consumers have acquired additional usage experience with the new product.

Trial-Repeat Model

The steady-state market share a new brand finally achieves can be represented directly as the product of the long-run levels of trial and repeat purchasing it attains. Following Parfitt and Collins [52], one can express market share for the new brand, $M(t)$, by:

$$M(t) = TS, \tag{8.13}$$

where: T = ultimate cumulative trial rate for the new brand, t (proportion of all buyers in the target group who ever try the new brand),

S = ultimate repeat purchase rate for the new brand, t (new brand's share of subsequent purchases in the product category made by buyers who have ever made a trial purchase of the new brand).

This model has been used extensively [2, 52] to forecast equilibrium shares ($M(t)$) for new brands using extrapolations of early test market measurements to estimate the ultimate trial (T) and repeat purchase (S) rates of equation 8.13. Here the authors employ a model previously used by Urban [67] which decomposes these two quantities slightly. The purpose is to represent the influence of certain marketing policy variables on consumer response in a simple fashion and at the same time make use of measurements obtained from the laboratory and post-usage studies.

One assumes that trial comes about in one of two ways: (1) receipt and use of free samples or (2) initial purchases. The incidence of first purchase of the new brand is taken to be dependent on the level of awareness induced by advertising or other forms of promotion and the extent of its retail availability. As an approximation, the probability of becoming aware of the new brand and that of having it available are presumed to be independent. One also assumes that the probability a consumer makes a first purchase is independent of the probability of receipt and use of a sample. Putting these assumptions together, one can model trial by:

$$T = FKD + CU - (FKD)(CU), \tag{8.14}$$

where: F = long-run probability of a consumer making a first purchase of the new brand given awareness and availability of it (i.e., proportion of consumers making a trial purchase in the long run given that all consumers were aware of it and distribution was complete),

D = long-run probability that the new brand is available to a consumer (e.g., proportion of retail outlets that will ultimately carry the new brand weighted by their sales volume in the product category),

K = long-run probability that a consumer becomes aware of the new brand,

C = probability that a consumer will receive a sample of the new brand,

U = probability that a consumer who receives a sample of the new brand will use it.

The various probabilities are averages for the particular group under consideration. As an estimator of F, one uses the proportion of respondents who purchased the new brand (Q_4 in table 8-1) in the laboratory on their simulated shopping trip. The next three parameters, K, D, and C, depend on the type and magnitude of marketing effort management plans to use if

the brand is test-marketed or otherwise launched. Thus, a prime determinant of the level of awareness (K) for the new brand is the amount to be spent for media advertising, whereas the extent of availability (D) depends on how much salesforce and promotional activity will be directed at the retail trade. The translation of the introductory marketing plan into estimates of K and D is accomplished by informal means, drawing upon managerial judgment as well as results and experience obtained with similar products. Analyses of certain types of historical data also can be helpful as, for example, in formulating a relationship between brand awareness and media expenditures or coverage. Estimation of the sampling coverage parameter (C) is straightforward, given knowledge of the scale of sampling program planned. Previous research with similar products or a small experiment can be used to estimate sample usage (U).

Urban [67] models the other quantity in equation 8.13, S, as the equilibrium share of a first-order, two-state Markov process:

$$S = \frac{R(k,t)}{1 + R(k,t) - R(t,t)}, \tag{8.15}$$

where the transition probabilities are defined as follows.

$R(k,t)$ = probability that a consumer who last purchased *any* of the established brands (k) will switch to the new brand (t) on the next buying occasion,

$R(t,t)$ = probability that a consumer who last purchased the new brand will repurchase it on the next buying occasion.

Estimates of $R(k,t)$ and $R(t,t)$ are derived from measurements obtained in the post-usage survey (Q_5 in table 8.1). The proportion of respondents who make a mail order repurchase of the new brand when given the opportunity to do so is taken as an estimate of $R(t,t)$. To estimate $R(k,t)$ for those who do *not* repurchase the new brand in this situation one uses their preference measurements for the new and relevant established brands obtained from them in the post-usage survey. Probabilities of purchasing the new brand are computed for each such individual by equation 8.4 and their average value is taken as an estimator of $R(k,t)$.

It is sometimes observed empirically that respondents who "purchased" the new brand in the laboratory experiment differ from those who received it as a free sample with respect to their repeat rates, S. In this case, separate repeat rates are calculated and applied to the appropriate trial components in equation 8.14 to adjust for the difference.

Applying the inputs to equations 8.14 and 8.15 gives estimates of the ultimate trial (T) and repeat (S) rates, respectively, which are then simply

multiplied together as indicated by equation 8.13 to calculate the expected long-run market share for the new brand.

This trial-repeat model is clearly a very simplified representation of the new product response process. Some tests of the adequacy of the model's overall structure have been reported by Urban [67]. He derived the various inputs required by the trial and repeat equations for several new products from studies of their test markets or national introductions. He then compared the ultimate trial and repeat rates and equilibrium market shares predicted by the model with the values of these quantities that actually had been observed. For each of the half dozen cases examined, he found the observed and predicted values to be in very close agreement.

In terms of its complexity, the foregoing model has proved to be adequate for the level of detail typically specified in introductory marketing plans at the stage of a new brand's development where the decision to test market or not is under consideration. An important assumption implicit in the model is that the frequency of purchase of the new brand will be the same as that for established brands. This assumption can be relaxed somewhat by weighting the ultimate repeat rate (S) by an index that reflects the new brand's usage rate in relation to that for established brands [52, 67]. Clearly the latter is at best a crude adjustment and situations can arise where, if the required measures can be obtained, it will be desirable to use one of the available models (e.g. 66] that allows the adoption process to be represented in greater detail.

Structural and Output Comparisons

The expression for market share developed from the individual preference-purchase probability model (equation 8.5) is structurally equivalent to that defined in terms of trial and repeat purchase levels (equation 8.13). In the former case, market share is the product of the relevant set proportion ($E(t)$) and the average conditional probability of purchasing the new brand ($\sum_{i=1}^{N} L_i(t)/N$). In the latter case, market share is the product of the cumulative trial proportion (T) and the share which repeat purchases of the new brand represent of subsequent buying by previous triers (S).

Though not precisely identical, "relevant set" and "trial" are operationally very similar constructs in the present context. As noted in the discussion of measurement procedures, the composition of a consumer's relevant set is determined by responses to a series of questions about which brands he/she has ever *used* or would consider using or not using. Thus, one would expect to find that brands so evoked for the most part are accounted for by past usage or "trial," and empirically this is the case. For example, in separate studies of three different product classes, 90% or more

of all brands respondents deemed relevant were identified on the basis of usage-related questions.

The quantities $\sum L_i(t)/N$ and S are both average conditional probabilities or shares of repeat purchases. However, they are distinguished conceptually in that the former are obtained from a zero-order individual level model whereas the latter arise from an aggregate first-order Markov process. Despite these differences, it is often difficult to distinguish between these two types of models, each of which may yield satisfactory results [45]. For example, aggregation over heterogeneous consumers will tend to *overestimate* the true order of the process [45, chapters 3 and 4]. However, Kesavan and Srinivasan [40] have shown that aggregation of several brands into a single "other" brand category (as is done here) will tend to *underestimate* the true order of the process and can lead to biased steady-state market share predictions. A second difference is that the average purchase probability obtained from the preference model ordinarily will reflect some effects of in-store promotion whereas these are not incorporated explicitly in the estimate of the repeat rate. The parameter of the preference-purchase probability model is estimated from data pertaining to the purchase of established brands which are supported by some level of in-store promotional activity, but no provision for such effect is made in the repeat submodel (equation 8.15).

The submodels and measures used to arrive at estimates of these conceptually similar quantities are distinct. Whereas the trial and repeat proportions are based on essentially *direct* observations of these quantities obtained under controlled conditions, the relevant set proportion and the average conditional purchase probability are estimated *indirectly* from other measures. Coming from the same research design, the measurement inputs for both models are affected by common sources of methods variance. Nonetheless, because of differences in the submodels and their respective inputs, agreement between the two market share predictions is not a built-in or guaranteed feature of these approaches and hence it is possible to make a meaningful check for convergence here.

Finding that the two models do yield outputs that are in close agreement can serve to strengthen confidence in the prediction. In contrast, divergent forecasts trigger a search for and evaluation of possible sources of error or bias that might account for the discrepancy. The first step is to compare the relevant set proportion ($E(t)$) and trial (T) estimates. Lack of agreement here could imply that the assumptions concerning awareness (K) and retail availability (D) are not compatible with those made implicitly or explicitly in estimating the relevant set proportion ($E(t)$) as, for example, when the latter is based on a regression of relevant set proportions on awareness levels for established brands. Given that these assumptions did appear compatible, then the possibility of measurement bias in the conditional trial probability (T) would be investigated.

After reconciliation of the trial and relevant set estimates, attention is focused on the values of the conditional purchase probability and the repeat rate. In comparing these quantities, it is important to keep in mind that effects of in-store promotional support are not represented in the repeat rate estimate derived from the post-usage interview. For product classes having substantial in-store promotional programs, upward adjustments in these initial estimates of repeat rates are necessary and justifiable. In the end, some judgment may have to be exercised to reconcile differences that arise, but that process is facilitated by careful consideration of the structural comparability of the two models.

Predictions and Marketing Plans

Prediction of a new brand's market share must reflect plans for the marketing program to be used in the future test market or launch. Frequently at this pre-test-market stage management is interested in evaluating some variations in the introductory marketing mix for the new brand. The trial-repeat model can be used to advantage in performing some rough-and-ready simulations of the effects of certain kinds of marketing mix modifications. Some of the changes or alternatives management may wish to consider can be approximated by judgmentally altering parameter levels. For example, increasing the level of advertising spending could be represented by raising the awareness probability K in equation 8.14. Differences in sampling programs could be handled similarly by modifying the C and U parameters. Other types of changes, as in advertising copy or price, that affect the conditional first purchase probability F, can be measured by expanding the research design shown in table 8-1 to allow observation of the differential effects on trial purchases made in the controlled shopping environment for alternative price or copy treatments.

After examination of the impact of strategic changes, profitability measures can be calculated for the market share estimates. On the basis of these inputs, management must decide whether or not to test market the new brand.

Estimation

At several points in the preceding discussion, reference is made to how data obtained from the laboratory and post-usage phases of the consumer research could be related to the models' parameters and input requirements. For the most part, this is a straightforward task involving only simple computations. However, estimation of the preference scale values, the para-

meter of the purchase probability model, and the relevant set proportion is somewhat more complex and is discussed in detail hereafter.

Preference Scaling

Data obtained by the constant sum, paired comparison method are used to estimate a vector of brand preference scale values for each respondent by the least-squares procedure proposed by Torgerson [65, pp. 109–12]. Respondents' preferences are scaled twice, before and after using the new brand. The "before" scaling is carried out with reference to respondent's idiosyncratic relevant sets identified by the premeasurement (Q_2 in table 8-1) and the "after" scaling (Q_5) encompasses the previously determined relevant set of established brands plus the new brand.

Under the assumption that the comparative judgments reported by a subject for stimuli reflect the ratios of their corresponding subjective magnitudes, then the least squares estimate of the stimulus scale has ratio scale properties. That the computed estimates actually attain this level of measurement cannot be verified from the input data and no statistical test for goodness of fit is available. Two types of internal consistency checks which bear on the quality of the preference scale estimates have been performed with data for deodorant and antacid categories obtained from separate samples. First, very few instances of intransitivities in preferences were uncovered when Kendall's method of circular triads was applied to each subject's paired comparison judgments [39, chapter 11]. The absence of inconsistencies is not a very demanding requirement here inasmuch as transitivity is only a necessary condition for the existence of an ordinal scale [23, p. 13–24] and with typical relevant set sizes of three to five brands, the number of paired comparison judgments required of subjects is most often small. As Torgerson [65, p. 116] suggests, a goodness-of-fit check also was made. The matrix of ratios representing a respondent's original paired comparison judgments was compared with the equivalent matrix calculated from the estimated brand preference scale values for that respondent. For the great majority of respondents, the estimated scale values for an individual's relevant set of m brands could very accurately reproduce the $m(m - 1)/2$ observed ratios that person had provided in performing the paired comparison judgments [5].

Estimation of the Purchase Probability Function Parameter

In the pre-exposure interview, the brands last purchased by respondents are identified and preference measures are obtained for their sets or relevant

alternatives. This information on last brand purchase and the brand preference scale values are used to estimate the parameter β of the purchase probability model defined in equation 8.2. Recall from the discussion of the preference model that β is to be estimated across different (established) brands and across respondents. Also, the observations are (dichotomous) purchase events, not probabilities. Now because

$$[\hat{V}_i(j)]^{\beta} = \exp[\beta \ln \hat{V}_i(j)],$$

one can write the purchase probability model (equation 8.2) as:

$$P_i(j) = \frac{\exp[\beta \ln \hat{V}_i(j)]}{\sum_{k=1}^{m_i} \exp[\beta \ln \hat{V}_i(k)]}. \tag{8.16}$$

The form of the expression is that of the multinomial logit model which McFadden [46] derived as a theory of population choice behavior. Maximum likelihood estimation procedures have been developed for this model and McFadden notes that the estimators obtained are asymptotically efficient and normally distributed under "very general conditions" [46, p. 119].[2] This method has been applied widely in economic studies of choice behavior [24] and is used here to estimate the β parameter in equation 8.16. More specifically, a program developed by Manski and Ben Akiva [13] is employed which uses the Newton-Raphson iterative technique to determine the value of the parameter β which maximizes the following likelihood function:

$$L = \prod_{i=1}^{N} \prod_{k=1}^{m_i} [P_i(k)]_{\delta ik}, \tag{8.17}$$

where $\delta_{ik} \begin{cases} = 1 \text{ if individual } i \text{ last purchased brand } k. \\ = 0 \text{ otherwise.} \end{cases}$

Though standard errors and associated t statistics for the β parameters in equation 8.16 can be obtained, the usual goodness-of-fit measure, the coefficient of determination (R^2), cannot be applied here because the estimated equation predicts probabilities whereas the observed values are purchase events (0, 1 measures). However, Hauser [35, chapters 10, 36] recently developed useful measures for assessing the fit of this model based on information theory concepts. Hauser views the model (16) as an information system—i.e., the probabilities obtained from the preference model provide information about the choice outcomes. Now the *prior* entropy measures the total uncertainty in the system *before* observation of the

preference data. To compute the prior entropy, Hauser proposes that a naive model be assumed whereby every member of the sample is assigned a probability of purchasing any brand ($P'(k)$) equal to its aggregate market share among the total sample's reported last purchases. Under this assumption, he demonstrates that the prior entropy is given by:

$$Z = - \sum_{k=1}^{m^*} P'(k) \log P'(k), \tag{8.18}$$

where: Z = total uncertainty in the system with m^* alternative brands,

$P'(k)$ = prior probability of choice of brand k, $k = 1, \ldots, m^*$.

After application of the observed data to the preference model (8.16), the uncertainty is reduced to the *posterior* entropy. Hauser shows that the amount by which the preference data reduce the prior entropy is the *expected information, EI,* provided by the model which is:

$$EI = \sum_{i=1}^{N} \sum_{k=1}^{m_i} P_i(k) \log \frac{P_i(k)}{P_i'(k)}, \tag{8.19}$$

where the $P_i(k)$ are obtained from equation 8.16.

Noting that the prior entropy also can be taken as a measure of how well a perfect model would perform, Hauser proposes that the "usefulness" of the model (8.16) be assessed by comparing the expected information, *EI,* with the prior entropy, *Z.* Thus an index of the model's usefulness can be defined as the proportion of total uncertainty removed or "explained" by the model:

$$G = \frac{EI}{Z} \tag{8.20}$$

Finally, Hauser shows the *observed* or *empirical information, OI,* is:

$$OI = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{m_i} \delta_{ik} \log \frac{P_i(k)}{P_i'(k)} \tag{8.21}$$

where $\delta_{ik} \begin{cases} = 1 & \text{when respondent } i\text{'s last purchase was brand } k, \\ = 0 & \text{otherwise.} \end{cases}$

He argues that with a large sample, the observed information should be

close to its expected value, and thus the "accuracy" of the model can be assessed by comparing *OI* and *EI*.

Table 8–3 shows the results obtained when the maximum likelihood procedure was applied to the preference and last brand purchased data from the deodorant study. Note that the estimated value of β is nearly 10 times its estimated standard error and the model accounts for slightly more than three-quarters of the total uncertainty present as measured by the index *G*. As expected, the value of *OI* is very close to that of *EI*.

As another check on the adequacy of the fit obtained, the estimated value of β in equation 8.16 was used to calculate each individual's probability of having last purchased each brand in his/her relevant set. These probabilities were aggregated to calculate the fitted value of each brand's expected share of last purchases. The latter can be compared with the observed shares. Across all 18 brands the mean absolute deviation was found to be .8 of one market share point (percentage). This value can be compared with an average absolute deviation of 2.5 market share points obtained for a "naive" model whereby an individual has the same probability of purchasing any brand in his/her evoked set—i.e., $P_i(j) = 1/m_i$. Figure 8–2 shows a plot of the observed and fitted shares. The largest deviations were for the two major brands where the model overpredicted their shares by 2.0 and 3.1 share points, respectively.

In the present context, the "independence of irrelevant alternatives" assumption underlying the Luce-McFadden models implies that β should not vary with relevant set size. To investigate this matter, the model (8.16) was estimated separately within groups defined by relevant set size. Table 8–3 shows the results. Some variation in the estimated β can be seen. However, none of the four β estimates is significantly different from the overall or total sample value at the .05 level. Making all possible pairwise comparisons among the four values for the different relevant set sizes, one finds only two of the six differences to be significantly different at the .05

Table 8–3
Maximum-Likelihood Estimation Results

				Fit Indices		
	Sample Size	$\hat{\beta}$	*Standard Error*	*G*	*EI*	*OI*
Total sample	279	2.09	.20	0.77	1.72	1.64
By relevant set size						
Two brands	85	1.84	.41	0.83	1.56	1.53
Three brands	90	2.75	.49	0.84	1.74	1.61
Four brands	65	2.20	.37	0.72	1.66	1.56
Five or more brands	39	1.80	.36	0.55	1.23	1.12

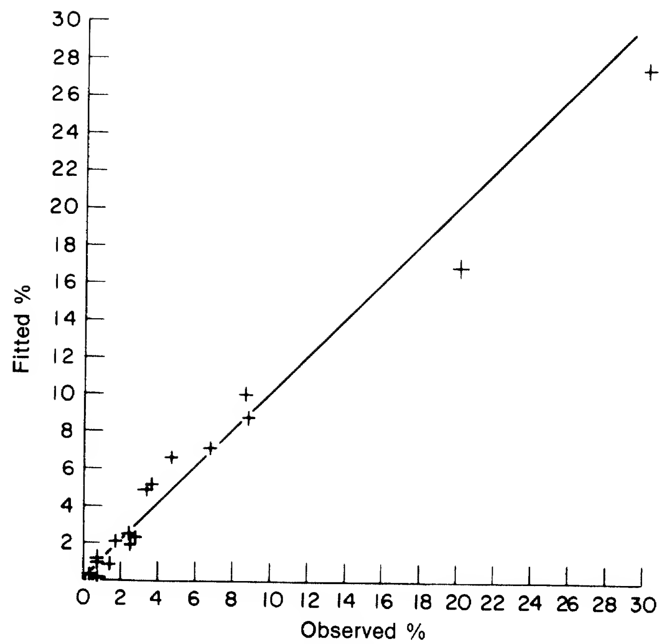


Figure 8–2. Plot of Observed versus Fitted Market Shares (N = 18 brands)

level. The quality of the fit as measured by the G index diminishes as the relevant set size increases but the sample sizes for the two largest relevant set size groups are also smaller.

As another test, equation 8.16 was estimated separately for each pair of brands within the subsample of respondents whose relevant set size was three. None of the β estimates so obtained differed significantly at the .10 level from the value obtained by estimating the parameter across all brands. These results do not appear to indicate any systematic contradiction of the assumption of independence of irrelevant alternatives for these data.

Estimation of the Relevant Set Proportion

Recall that from the preference model one obtains an estimate of the probability of purchasing the new brand that is conditional on its being a relevant choice alternative. Thus, one requires a method of predicting what proportion of consumers in the target group will eventually include the new brand in their relevant sets [$E(t)$ in equation 8.5].

In the discussion of the comparability of the trial-repeat and preference

models, it was noted that for the operational definition used here almost all the brands composing consumers' relevant sets are those with which they report having had some usage experience. Also there tends to be a strong and stable concurrent relationship across brands between aggregate levels of brand awareness and usage [e.g., 14]. Thus similar relationships between relevant set and awareness proportions are suggested, and have been found in the present work. To illustrate, cross-sectional regressions of relevant set proportions ($E(j)$) on unaided brand awareness ($B(j)$) and advertising awareness ($AA(j)$) levels were performed for 18 established brands of deodorants using measures of these variables obtained in the premeasurement questionnaire (Q_2 in table 8–1). Given that the observations were proportions which varied considerably in magnitude, an arcsin transformation was applied to them as a means of stabilizing the error variance and thereby obtaining efficient estimates from ordinary least squares regressions. The following results were obtained:[3]

$$\begin{aligned} \text{Arcsin}\, E(j) &= -.599 + \underset{(23.4)}{.901}\, \text{Arcsin}\, B(j) + e(j), && (8.22) \\ R^2 &= .972,\ \text{SEE} = 2.39. \\ \text{Arcsin}\, E(j) &= 3.91 + \underset{(11.6)}{1.066}\, \text{Arcsin}\, AA(j) + e(j), && (8.23) \\ R^2 &= .894,\ \text{SEE} = 4.61. \end{aligned}$$

As expected, both brand and advertising awareness appear to covary with the relevant set measure. However, the values of the coefficient of determination (R^2) and the standard error of estimate (SEE) indicate that the brand awareness regression provided a better fit of the data than did the estimated advertising awareness equation. Transforming the estimated values of the arcsin of $E(j)$ from the regression back to proportions and comparing them to their corresponding observed values, one finds the average residual for the brand awareness regression to be .021 and that for the advertising awareness regression to be .041.

To estimate the expected relevant set proportion for the new brand ($E(t)$), one simply applies the level of unaided brand awareness ($B(t)$) which the introductory marketing program is expected to achieve to the brand awareness equation. As noted in the discussion of the trial-repeat model, the level of brand awareness predicted for the new product is largely a judgmental estimate because it depends on the nature and magnitude of marketing effort that will be applied to support the introduction of the new brand.

Application

Background

The first application of the methodology involved a new brand of an aerosol deodorant product introduced by a competitor of the firm which sponsored the present work. Annual sales (at retail) for the product class in the United States amount to almost a half billion dollars and approximately a score of national brands were already being marketed before the emergence of the new brand. However, the two leading established brands held nearly half the market and the next five largest brands accounted for another 35% of the total product category volume. The new brand had been developed carefully and the basis of its positioning strategy was a straightforward but powerful claim of superior performance on an important attribute. The appearance of the new brand in a test market was regarded by management in this field as a competitive event of major importance.

Application of the system to this problem situation began after the new brand had been in test market in a midwestern city for eight months. Thus, carrying out the field work in a city *different* from that where the test market was underway afforded an opportunity to perform a test of the system's predictive ability in a relatively short period of time.

The design and conduct of the data collection corresponded to the methods and procedures summarized in table 8-1. Two hundred and ninety-nine respondents were interviewed in a suburban shopping center of a city separate from, but similar to, the site of the test market then in progress. Quota sampling was used to obtain the desired representation of demographic characteristics and usage habits among persons interviewed. Respondents were shown television commercials for the five leading established brands plus one for the new brand. After giving his or her reaction to the commercials on a small set of rating scales, each respondent entered the simulated store with a coupon worth $2.00 in cash. Prices were set to be equal to the average of those prevailing in discount stores in the area at that time. Almost 75% of the sample bought one or another of the brands available which included the new one. Those who did not purchase the new brand were given a free sample as they left the store. Post-usage interviews were conducted by telephone three weeks later. Because the product is one typically used daily, this period was long enough for respondents to accumulate considerable usage experience with the new brand. Two-thirds of the original sample were reinterviewed and had been using the new brand.

Results

At several points in the foregoing discussions of measurement and estimation, data from the deodorant study were used to illustrate methods and

Table 8-4
Model Inputs

	Quantity	*Initial Estimates*	*Adjusted Estimates*	*Revised after Test Market*
F	Conditional first purchase	.16	—	—
K	Brand awareness	.80	—	.83
D	Availability	.90	—	—
C	Sample coverage	.40	—	—
U	Sample use	.75	—	.67
$R(k,t)$	Repeat rate	.11	.20	—
$R(t,t)$		.41	.55	—
Trial repeat model				
T	Cumulative trial	.381	—	.356
S	Repeat share	.157	.308	—
$M(t)$	Predicted market share (percent)	6.0	11.7	11.0
Preference model				
$E(t)$	Relevant set proportion	.383–.445	—	.33
$\Sigma L_i(t)/N$	Avg. purchase prob.	.32	—	—
$M(t)$	Predicted market share (percent)	12.3–14.2	—	10.6

results. Here attention is focused on the main predictions obtained from the models. Table 8-4 summarizes the inputs for the two models.

From the preference model, the average post-trial purchase probability for the new brand ($\sum L_i(t)/N$) was estimated to be .32. Estimates for the relevant set proportion ($E(t)$) were obtained by translating the $10 million annual national advertising spending rate estimated for the new brand into expected levels of advertising and brand awareness and then using these values in the cross-sectional regression equations (8.22 and 8.23) for established brands to obtain predictions of the evoking proportion. This process yielded predictions for $E(t)$ of .383 (from the advertising awareness equation) and .445 (from the brand awareness equation). When combined in equation 8.5 with the estimate of .32 for the average post-trial purchase probability, these values of the relevant set parameter led to predicted market shares for the new brand of 12.3% and 14.2%, respectively.

The share prediction initially calculated from the trial-repeat model was much lower than the values obtained from the preference model. Although the value of the conditional probability of first purchase (F) estimated from the observed purchase rate of the new brand in the laboratory store was only .16, it was expected that a considerable amount of trial usage would be effected by a very extensive sampling program. When the introductory marketing plan was translated into the quantities specified in equation 8.14 it yielded a predicted value of .381 for the ultimate cumulative trial rate (T). This level of trial was consistent with the values of the evoking proportion (.38 to .445) estimated for the preference model. However, the repurchase

inputs derived from the post-usage survey when applied to equation 8.15 led to an estimate of only .157 for the ultimate repeat purchase rate (S). This repeat (S) estimate plus the trial (T) level .381 gave a predicted share of 6.0% for the new brand. The latter share was about half the 12.3–14.2% level predicted by the preference model, which used an average post-trial purchase probability estimate ($\sum L_i(t)/N = .32$) roughly double the magnitude of the ultimate repeat rate quantity ($S = .157$) estimated for the trial-repeat model.

Accounting for this marked discrepancy in the two repurchase estimates and hence the market share predictions was problematical. Efforts to uncover the source of the difficulty finally suggested a plausible diagnosis related to the measurements of the components of the repeat purchase rate, S. Recall from the discussion of the trial-repeat model that $R(t, t)$ is estimated by the proportion of respondents who make a "mail order purchase" of the new brand when given the opportunity to do so in the post-usage survey. In this initial application, the parameter $R(k, t)$ was estimated from responses to a buying intentions scale rather than in the manner described heretofore which subsequently was adopted. As a consequence of these procedures, the estimate of the overall repeat rate S did not reflect any influence in in-store promotion or other external sources of reinforcement. However, such effects are implicitly represented in the calibration of the preference-purchase probability model. Furthermore, there was reason to believe that repurchase intentions expressed immediately after rejecting an opportunity to make a mail order purchase of the brand might be understated because respondents wished to discourage further solicitations. The influence of in-store promotion was known to be very important in this product category generally and the manufacturer of the new brand in particular has a reputation for using in-store activities aggressively as a means of stimulating repeat purchasing. For these reasons an upward adjustment of the observed levels of repurchase intentions appeared justifiable and the two repeat probabilities were raised judgmentally: $R(k, t)$ from .11 to .20 and $R(t, t)$, from .42 to .55. These modifications changed the ultimate repeat rate (S) from .157 to .308 and thereby raised the market share predicted by the trial-repeat model to 11.7% which is very close to the lower end of the 12.3–14.2% range obtained from the preference model.

The share prediction finally presented to management was the midpoint in the 12.3–14.2% range predicted by the preference model or 13.3% —a reflection of the greater confidence placed in the results obtained from the preference model in comparison with the trial-repeat model in this situation. The share observed in the test market, 12 months after launch, was 10.4%. The prediction exercise was carried out by the model building team while the new brand was in test market but before their exposure to any specific feedback or measurements on its early performance. As explained in the discussion of the models, the levels of certain marketing mix control

variables that will persist in a test market must be specified in advance in order to develop predictions from the models. Here, the management group sponsoring this work had to supply these inputs for a competitor's brand rather than their own, and thus precise prior information was not available. In the course of reviewing the test market results, some significant differences were uncovered between the assumptions about the new brand's marketing plan that had been used in developing the predictions from the models and what actually had taken place in the test market. Taking account of the advertising and sampling programs which had, in fact, been used in the test market implied changes in the parameter estimates as indicated in the last column of table 8–4 and, as expected, would have improved the accuracy of the market share prediction generated by the preference model. Whereas the difference between the share initially predicted and that observed in the test market was $13.3 - 10.4 = 2.9$ share points, the "revised deviation" or difference between the revised *ex post* prediction and the observed share was only $10.6 - 10.4 = 0.2$ share point.

Discussion

The foregoing discussion of the first application illustrates how features of the system and understanding of its capabilities and limitations have evolved. As additional applications have taken place, the adaptability of the procedures has been tested and certain modifications have been introduced to cope with new problems and to effect improvements. In the first study, the preference and trial-repeat models produced very different market share predictions and judgment had to be exercised to reconcile the discrepancies and arrive at a final prediction. After this experience, the change in the method of estimating the $R(k,t)$ parameter was adopted, but a completely satisfying explanation of the discrepancy has never been found. The practice of using both models has been continued and in more than 30 subsequent applications differences of the magnitude of those that arose in the first study have not been encountered. A Monte Carlo analysis performed in the trial-repeat model gave an estimate of 1.6 share points for the standard deviation of the model's market share predictions. This figure provides a rough basis for assessing disparities in the predictions given by the two models. If the differences appear to be within the bounds of sampling fluctuations, a simple average of the two outputs is used as the share prediction. When more substantial discrepancies occur, they must be interpreted; and so judgment, guided by an examination of the diagnostic information obtained at several points in the measurement process outlined in table 8–1, ultimately plays a role in deciding which results should be relied upon to obtain a final share prediction.

Available external data are not sufficient to allow clear discrimination between the two models and their measurement inputs. The time lags, attrition, and other exigencies normally encountered in the development of new packaged goods have made for a slow accumulation of opportunities for acquiring validation information. No tightly controlled tests of the present system's predictive accuracy have been performed. Cases in which new products have been subjected to both ASSESSOR *and* test market evaluations provide a basis for an early but only partial assessment of the quality of predictions generated. Of the approximately 30 new packaged goods studied with ASSESSOR to date, test markting has been completed for nine and their final test market shares are known. Table 8-5 is a summary of results for these cases. The products are listed in the chronological order in which they were studied, beginning with the first application to the deodorant product which is included for completeness. Note that the ASSESSOR studies for the first three products were performed while their test markets were in progress and so are labeled "concurrent." These three applications were performed when the system first was developed and were conducted in this manner at the request of firms seeking information that would enable them to make an early evaluation of the system's predictive capability. In each of these nine cases, the ASSESSOR investigation was carried out in a city different from that used for the test market.

Table 8-5 shows the differences in the *initial* share predictions given by the preferences and trial-repeat models, *before* any reconciliation—i.e., the predictions based on planned or assumed test-market programs. Except in the first application to the deodorant product, the discrepancies did not exceed one share point. For all nine products, including the first, the absolute average deviation was 1.2 share points, which indicates that the predictions obtained from the two models generally have been in close agreement.

Also presented in table 8-5 are the observed test market shares and the final share predictions made *after* comparing and (where necessary) reconciling judgmentally the separate predictions derived from the two models, but *before* test market results were known. Hence, these predictions do *not* reflect any *ex post* adjustments made to account for differences between planned and actual or implemented levels of marketing effort used in the test markets.

As can be seen from the table, the deviations between the original predictions and the observed shares generally have been small, their absolute average being slightly less than one share point. However, the deviations in some instances appear more substantial when viewed as a percentage of the observed share, ranging from a low of 2% in the case of the fruit drink to a high of 50% for the pain reliever. As was noted for the deodorant application, seldom will the marketing mix program assumed in developing a prediction before the test market correspond exactly to that which is actu-

Table 8–5
Predicted and Observed Market Shares

Product	*Timing of Pretest in Relation to Test Market*		*Difference in Share Predictions of Preference and Trial-Repeat Models*[a]	*Market Share (percent)*		*Deviation*[b]
	Concurrent	*Before*		*Predicted*	*Observed*	
Deodorant	✓		+7.3	13.3	10.4	+2.9
Antacid	✓		−0.9	9.6	10.5	−0.9
Laundry ingredient	✓		+0.1	1.8	1.8[c] 2.0	−0.1
Household cleanser		✓	−0.4	12.0	12.5	−0.5
Shampoo		✓	+0.7	3.0	3.2	−0.2
Dishwashing ingredient		✓	−0.2	9.3	8.5	+0.8
Pain reliever		✓	+1.0	3.0	2.0	+1.0
Fruit drink		✓	−0.2	4.9	5.0	−0.1
Cereal		✓	+0.1	6.0	4.4	+1.6
Average (Absolute)			1.2	7.0	6.5	0.9

[a]Market-share prediction obtained from the preference model minus that obtained from the trial-repeat model.

[b]Predicted minus observed market shares.

[c]Shares observed in two test-market cities. The "observed" share used to calculate the "deviation" for this product was the mean of these two figures.

ally implemented later. Not surprisingly then, it has also been found for several of the subsequent applications that *ex post* predictions based on more precise knowledge of the marketing efforts expended in the test markets deviate less from the observed shares than do the original predicttions shown in table 8–5.

These results are reported in the spirit of revealing what is known about the accuracy of predictions developed through use of the system, but clearly they do not constitute a true predictive test. Though all the applications completed to date for which test market shares are available have been included, these cases are few in number and did not arise in a planned or prespecified manner. The lack of uniformity and precision associated with the observed test market shares themselves also deserves emphasis. The figures were obtained from several firms for the particular products whose investigation they had sponsored. Thus, the observed shares originated from several different sources, using a variety of methods. No claim can be made that the conditions of equivalence and independence have been met that enable unequivocal inferences about external validity to be drawn from comparisons of predicted and observed events.

The adequacy of the model's predictive ability must be evaluated in relation to how the model is used. At the pre-test-market stage, the manager is most interested in knowing whether he has a "winner." Will the brand earn a substantial share of the market? The second issue of concern to managers is how to improve the product's performance. The system's diagnostic capabilities and ability to make conditional forecasts for strategic changes can aid the manager in this task. Finally, the manager wants to know whether to drop the product, go to test market, or go national. If the predicted share is low and feasible changes in the marketing plan do not have potential to improve share substantially, dropping the product would be appropriate. If the share is good, either going to test or national introduction would be possible. The model proposed here does not answer this question. The manager could consider going national if the share is very high, investment is small, and competitive imitation is a danger. Usually, the product would go to test market. However, the test market would be oriented toward finding improvements in the marketing strategy rather than determining whether the product can attain an adequate market share. In this environment, the test market would be designed to place emphasis on measurement of response to marketing variables rather than determining share. The use of test market analysis models [e.g., 66] would be appropriate to process such data. If the test market confirms pre-test share estimates, the product could be introduced.

Conclusions

The authors have described a set of models and measurement procedures intended for use in evaluating new packaged goods at that stage in their development where management is faced with the decision of whether or not to place them in test markets. The approach taken to this problem is to merge relevant behavior and management science concepts and methods. The results obtained from the initial applications have been sufficiently encouraging to suggest that the kind of methodology discussed can be a useful addition to the growing body of decision-support technology now available and being applied to the problems of managing new product development in the packaged goods field.

The system described is intended to aid management in evaluating a new packaged good brand at a particular point in the developmental process and it is important to recognize where the system can be expected to prove useful and where it may not. Experience gained from applications of the system as well as the nature of the models and measurement methodology itself suggest at least three factors or conditions as being necessary for obtaining satisfactory results. First, the applicability of the system is limited

to situations in which the new brand seeks to penetrate a product category well defined in terms of the nature and closeness of substitutes. Cases in which a very novel or innovative offering effectively creates a new product category cannot be handled by these methods. Second, the assumption that the usage/purchase rate for the new brand will be the same as that for the established brands must be tenable. Only limited means of coping with departures from this condition are available. A third restriction is that consumption and learning must occur at rates such that preferences for the new brand stabilize in a relatively short period. For products which are used infrequently or which require long periods of usage before benefits/satisfaction can be realized, it would not be feasible to measure post-usage preferences by the means described.

The development and evaluation of the system are an ongoing process. Additional tests bearing on the general issue of predictive validity will be possible in the future as test market data accumulate for products previously evaluated by this methodology. Future work will be undertaken to extend the range of new product situations to which the system can be applied.

Notes

1. The term *relevant set* is due to Allaire [3] and is akin to Howard and Sheth's [37] concept of "evoked set". The former consists of familiar alternatives, irrespective of how favorably (or unfavorably) they are evaluated. In contrast, evoked set generally has been interpreted to include only "acceptable" alternatives. For further discussion of this distincion as well as other conceptual refinements and operational definitions different from that used here, see Refs. 38 and 48.

2. The small sample properties of the maximum likelihood estimator of the multinomial logit model are, in general, unknown. However, on the basis of examples and Monte Carlo studies McFadden suggests that the approximation is "reasonably good." See the discussion in Ref. 46, p. 119 ff.

3. Both regressions are based on 18 observations. R^2 and SEE denote the coefficient of determination and the standard error of estimate, respectively. The figures in parentheses are the t statistics for the regression coefficients.

References

1. Achenbaum, Alvin A. "Market Testing: Using the Marketplace as a Laboratory," *in* Robert Ferber, ed., *Handbook of Marketing Research.* New York: McGraw-Hill Book Company, 1974, 4-31-54.

2. Ahl, David H. "New Product Forecasting Using Consumer Panels," *Journal of Marketing Research,* 7 (May 1970), 159-67.
3. Allaire, Yvan. "The Measurement of Heterogeneous Semantic, Perceptual, and Preference Structures," unpublished Ph.D. thesis, Massachusetts Institute of Technology, 1973.
4. Allaire, Y., Gilles Laurent, and Alvin J. Silk. "The Goodness of Fit of Constant Sum Preference Scaling," in preparation.
5. Anonymous. "The Rebuilding Job at General Foods," *Business Week,* No. 2294 (August 25, 1973), 48-55.
6. Assmus, Gert. "NEWPROD: The Design and Implementation of a New Product Model," *Journal of Marketing,* 39 (January 1975), 16-23.
7. Axelrod, Joel N. "Attitude Measures that Predict Purchase," *Journal of Advertising Research,* 8(March 1968), 3-18.
8. Ayal, Igal. "Trial and Repeat Buying—The Case for Separate Consideration," *Journal of the Academy of Marketing Science,* 2 (Summer 1974), 413-31.
9. Bass, Frank M. "The Theory of Stochastic Preference and Brand Switching," *Journal of Marketing Research,* 11 (February 1974), 1-20.
10. ———. "Analytical Approaches in the Study of Consumer Purchase Behavior and Brand Choice," Paper No. 520, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, July 1975.
11. Bass, F.M., Edgar A. Pessemier, and Donald R. Lehmann. "An Experimental Study of Relationships between Attitudes, Brand Preference, and Choice." *Behavioral Science,* 17 (November 1972), 532-41.
12. Bass, F.M., Abel Jeuland, and Gordon P. Wright. "Equilibrium Stochastic Choice and Market Penetration Theories: Derivations and Comparisons," *Management Science,* 22 (June 1976), 1051-63.
13. Ben-Akiva, Moshe. "Program for Maximum Likelihood Estimation of the Multinomial Logit Model," unpublished paper, Transportation Systems Division, Department of Civil Engineering, Massachusetts Institute of Technology, January 1973.
14. Bird, Michael, and A.S.C. Ehrenberg, "Non-Awareness and Non-Usage," *Journal of Advertising Research,* 6 (December 1966), 4-8.
15. Burger, Philip C. "COMP: A New Product Forecasting System," Working Paper 123-72, Graduate School of Management, Northwestern University, 1972.
16. Buzzell, Robert D., and Robert E.M. Nourse. *Product Innovation in Food Processing, 1954-1964.* Boston, Massachusetts: Division of

Research, Graduate School of Business Administration, Harvard University, 1967.

17. Cadbury, N.D. "When, Where, and How to Test Market," *Harvard Business Review,* 53 (May–June 1975), 96–105.
18. Campbell, Brian M. "The Existence of Evoked Set and Determinants of Its Magnitude in Brand Choice Behavior," unpublished Ph.D. dissertation, Columbia University, 1969.
19. Campbell, Donald T., and Julian C. Stanley. *Experimental and Quasi-Experimental Designs in Research.* Chicago: Rand-McNally and Company, 1963.
20. Charlton, P., A.S.C. Ehrenberg, and B. Pymont. "Buyer Behavior under Mini Test Conditions," *Journal of the Market Research Society,* 14 (July 1972), 171–83.
21. Charlton. P., and B. Pymont. "Evaluating Marketing Alternatives," *Journal of the Market Research Society,* 17 (April 1975) 90–103.
22. Claycamp, Henry J., and Lucien E. Liddy. "Prediction of New Product Performance: An Analytical Approach," *Journal of Marketing Research,* 6 (November 1969), 414–20.
23. Coombs, Clyde H., Robyn M. Dawes, and Amos Tversky. *Mathematical Psychology.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.
24. Domencich, Thomas A., and Daniel McFadden. *Urban Travel Demand.* Amsterdam: North-Holland Publishing Co., 1975.
25. Ehrenberg, A.S.C. *Repeat-Buying.* Amsterdam: North-Holland Publishing Co., 1972.
26. ———. "Predicting the Performance of New Brands," *Journal of Advertising Research,* 11 (December 1971), 3–10.
27. Ehrenberg, A.S.C., and G.J. Goodhardt. "Repeat Buying of a New Brand," *British Journal of Marketing,* 2 (Autumn 1968), 200–5.
28. Eskin, Gerald J., and John Malec. "A Model for Estimating Sales Potential prior to the Test Market," in *Proceedings* 1976 Fall Educators' Conference, Series No. 39. Chicago: American Marketing Association, 1976, 230–3.
29. Forecast Market Research, Ltd. *The Mini-Test Market.* London: 1976.
30. Ginter, James L. "An Experimental Investigation of Attitude Change and Choice of a New Brand," *Journal of Marketing Research,* 11 (February 1974), 30–40.
31. Green, Paul E. "Marketing Applications of MDS: Assessment and Outlook," *Journal of Marketing,* 39 (January 1975), 24–31.
32. Haley, Russell I. "We Shot an Arrowhead (#9) into the Air," *Proceedings,* 16th Annual Conference, Advertising Research Foundation, New York, 1970, 25–30.

33. Hardin, D.K. "Changing Test Market Technology in the U.S.A.," in *Marketing in a Changing World: The Role of Market Research,* ESOMAR Proceedings, 1972 Congress, Cannes, France. Amsterdam, The Netherlands: European Society of Opinion and Marketing Research, 315–39.
34. Hatch, Thomas E., and Glen L. Urban. "New Products—Minimizing Risk and Maximizing Creativity," Working Paper 752–74, Sloan School of Management, Massachusetts Institute of Technology, November 1974.
35. Hauser, John R. "A Normative Methodology for Predicting Consumer Response to Design Decisions: Issues, Models, Theory and Use," unpublished Ph.D. thesis, Massachusetts Institute of Technology, 1975.
36. ———. "The Information Measure Provides a Mathematically Rigorous and Intuitively Appealing Predictive Test for Probabilistic Demand Models," Working Paper, Transportation Center, Northwestern University, March 1976.
37. Howard, John A., and Jagdish N. Sheth. *The Theory of Buyer Behavior.* New York: John Wiley & Sons, Inc., 1969.
38. Jacoby, Jacob. "A Model of Multi-Brand Loyalty." *Journal of Advertising Research,* 11 (June 1971), 25–31.
39. Kendall, Maurice G. *Rank Correlation Methods.* London: Griffin, 1955.
40. Kesavan, R., and V. Srinivasan. "Bias in Stochastic Binary Models of Brand Choice," in *Proceedings* 1976 Fall Educators' Conference, Series No. 39, Chicago: American Marketing Association, 115–9.
41. Kotler, Philip. *Marketing Decision Making.* New York: Holt, Rinehart and Winston, 1971.
42. Lipstein, Benjamin. "The Design of Test Marketing Experiments," *Journal of Advertising Research,* 5 (December 1965) 2–7.
43. Luce, R. Duncan. *Individual Choice Behavior.* New York: John Wiley & Sons, Inc., 1969.
44. Massy, William R., Ronald E. Frank, and Thomas M. Lodahl. *Purchasing Behavior and Personal Attributes.* Philadelphia: University of Pennsylvania Press, 1968.
45. Massy, W.F., David B. Montgomery, and Donald F. Morrison. *Stochastic Models of Buyer Behavior.* Cambridge, Massachusetts: M.I.T. Press, 1970.
46. McFadden, Daniel. "Conditional Logit Analysis of Qualitative Choice Behavior," *in* Paul Zarembka, ed., *Frontiers in Econometrics.* New York: Academic Press, 1970, 105–42.
47. McGuire, E. Patrick. *Evaluating New Product Proposals.* New York: The Conference Board, Report No. 604, 1973.
48. Narayana, Chem L., and Rom J. Markin. "Consumer Behavior and

Product Performance: An Alternative Conceptualization," *Journal of Marketing,* 39 (October 1975), 1-6.

49. National Industrial Conference Board. *Market Testing Consumer Products.* New York: National Industrial Conference Board, Experiences in Marketing Management, No. 12, 1967.
50. Nielsen Marketing Service. "New Brand or Superbrand?" *The Nielsen Researcher,* 29, No. 5(1971), 4-10.
51. ———. "To Test or Not to Test," *The Nielsen Researcher,* 30, No. 4(1972), 3-8.
52. Parfitt, J.H., and B.J.K. Collins. "Use of Consumer Panels for Brand Share Prediction," *Journal of Marketing Research,* 5 (May 1968), 131-46.
53. Pessemier, Edgar A. "Market Structure Analysis of New Product and Market Opportunities," *Journal of Contemporary Business,* 4 (Spring 1975), 35-65.
54. Pessemier, E.A., and Richard D. Teach. "A Single Subject Scaling Model Using Judged Distances between Pairs of Stimuli," Paper No. 143, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, August 1966.
55. Pessemier, E.A., Philip Burger, Richard Teach, and Douglas Tigert. "Using Laboratory Brand Preference Scales to Predict Consumer Brand Purchases," *Management Science,* 17 (February 1971), B-371-385.
56. Pessemier, E.A., and William L. Wilkie. "Multi-Attribute Choice Theory—A Review and Analysis," Paper No. 372, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, September 1972.
57. Pymont, B.C. "The Development and Application of a New Micro Market Testing Technique," in *The Practical Applications of Market Research,* ESOMAR Papers, 1970 Congress, Barcelona, Spain. Amsterdam, The Netherlands: European Society of Opinion and Marketing Research, 201-21.
58. Pymont, B.C., D. Reag, and P.G.M. Standen. "Towards the Elimination of Risk from Investment in New Products: Experience with Micro-Market Testing," Paper presented at the 1976 ESOMAR Congress, Venice, Italy, September 1976.
59. Rao, Vithala. "The Salience of Price in the Perception and Evaluation of Product Quality: A Multidimensional Measurement Model and Experimental Test," unpublished Ph.D. thesis, University of Pennsylvania, 1970.
60. Shocker, Allan D., and V. Srinivasan. "A Consumer-Based Method-

ology for the Identification of New Product Ideas," *Management Science,* 20 (February 1974), 921-37.

61. Stanton, Frank. "What Is Wrong with Test Marketing?" *Journal of Marketing,* 31 (April 1967), 43-7.
62. Stefflre, Volney, "Market Structure Studies: New Products for Old Markets and New Markets (Foreign) for Old Products," *in* Frank M. Bass, Charles W. King, and Edgar A. Pessemier, eds., *Applications of the Sciences in Marketing.* New York: John Wiley & Sons, Inc., 1968, 251-68.
63. Stevens, S.S. "Ratio Scales of Opinion," *in* Dean Whitla, ed., *Handbook of Measurement and Assessment in Behavioral Sciences.* New York: Addison-Wesley Publishing Company, 1968, 171-99.
64. Tauber, Edward M. "Forecasting Sales Prior to Test Market," *Journal of Marketing,* 41 (January 1977), 80-4.
65. Torgerson, Warren S. *Theory and Method of Scaling.* New York: John Wiley & Sons, Inc., 1958.
66. Urban, Glen L. "SPRINTER Mod III: A Model for Analysis of New Frequently Purchased Consumer Products," *Operations Research,* 18 (September-October 1970), 805-53.
67. ———. "PERCEPTOR: A Model for Product Positioning," *Management Science,* 21 (April 1975), 858-71.
68. Urban, Glen L., and Richard Karash. "Evolutionary Model Building," *Journal of Marketing Research,* 8 (February 1971), 62-6.
69. Wachsler, Robert A., Lewis G. Pringle, and Edward I. Brody. "NEWS: A Systematic Methodology for Diagnosing New Product Marketing Plans and Developing Actionable Recommendations," unpublished paper, Research Department, Batten, Barton, Durstine and Osborn, Inc., New York, February 1972.
70. Wierenga, B. *An Investigation of Brand Choice Processes.* Rotterdam: University of Rotterdam Press, 1974.
71. Wilkie, William L., and Edgar A. Pessemier. "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Research,* 10 (November 1973), 428-41.
72. Wind, Yoram. "A New Procedure for Concept Testing," *Journal of Marketing,* 37 (October 1973), 2-11.
73. Winter, Frederick W. "A Laboratory Experiment of Individual Attitude Response to Advertising Exposure," *Journal of Marketing Research,* 10 (May 1973), 130-40.

9 LTM Estimating Procedures

Yankelovich, Skelly and White, Inc.

This chapter details the basis for and procedures used in computing market-share and volume estimates in the LTM simulation system (LTM = Laboratory Test Market, a research service of Yankelovich, Skelly and White, Inc.).

The description focuses on the *basic* LTM experiment, operative since 1968, after five prior years of development. The LTM *now* offers four additional experimental models, each dependent on a somewhat different estimating model; these, too, are covered in this chapter. *The first section* presents the framework for LTM estimates. *The second section* describes the estimating process in general and for each of the five LTM models.

The Framework for LTM Estimates

An LTM study is essentially a laboratory experiment incorporating all of the factors which affect product acceptance by the consuming public, and including the opportunity to purchase by consumers with their own money in a laboratory store.

Of course, the laboratory situation cannot be fully identical to the real world. The LTM estimating procedure therefore entails a translation of the raw experimental data to real world equivalents. This section explains how this translation is done by describing: (1) collection of data used in making estimates, and (2) differences between the laboratory and the real world.

Collection of Data Used in Making Estimates

In essence, the Laboratory Test Market system provides an opportunity for each key element of a real world new product introduction to act upon the consumer—but under controlled conditions, and with substantial compression of the process in both time and space. These key elements are:

A mechanism for generating awareness of the new product.

An internal publication of Laboratory Test Market, a division of Yankelovich, Skelly and White, Inc., New York, N.Y. Reprinted with permission.

A means of tapping into trial interest—if it is genuinely present.

An opportunity for the consumer to "buy," with some "investment" by the participant (most desirably, the spending of money).

An opportunity to collect vivid, almost immediate feedback on: the factors which encouraged and/or retarded buying of the new product; expectations about the new product on the part of buyers; the degree to which consumers were confused, vague, or just plain wrong about the new product—even after exposure to its planned communication, support, etc.

An opportunity for normal in-house use of the new product by buyers, without any knowledge that they are still participating in an experiment.

A mechanism for gathering information on how the new product was used; how often; what products, if any, it displaced; how it performed against expectations.

A measure of repeat/conversion rates among buyers.

A measure of whether frequency of use and/or purchase of the new product will follow norms or will be higher or lower than other products in the category.

Further, the Laboratory Test Market system allows the above eight elements to operate in the context of the current competitive climate.

The need to allow the above eight elements of a new product introduction to operate under "laboratory" conditions governed the design of the Laboratory Test Market facilities which are used. Specifically, while the physical locations at which Laboratory Test Market studies are carried out may vary in looks and decor, they have in common:

A theater-like area, where consumers can be exposed to TV commercials, slides, print ads, etc., i.e., to the essence of the communication themes and executions for the new product and for key competition. The theater provides the setting for the staging of media activity; it is the place where awareness and trial interest come into play.

A store replica, where—along with relevant competitive products—the consumer can see and examine the new product in a typical display (with pricing, and any planned in-store supports, etc.) The "store" is the place where a "purchase" can be made of the new product or of a competitive product, or where the decision to buy nothing can be reached. Although respondents know they are not in a real store, the facility is designed to look sufficiently authentic to encourage them to move into a "shopping mode."

A check-out arrangement where consumers can pay for and collect their purchases before they leave the Laboratory.

Areas where consumers can report on what they did and did not do in the store and why.

A capability to talk with the consumer (via telephone, for example) at a reasonable interval after the purchase—say after two to six weeks. (The time interval is contingent upon the frequency with which the particular product class is used.) This capability provides the basis for gathering of the necessary usage and product-reaction data, of the repeat conversion rates, and of the critical frequency-of-use data.

A capability to remeasure purchase and usage data at several intervals after initial exposure in the Laboratory (sales waves)—particularly relevant to product categories where "wear-out" may occur.

Experience in applying the Laboratory Test Market system over the past eleven years suggests that it is important that all the different areas within the physical location be separated from one another so that participating consumers cannot anticipate the next step in the procedure. Further, a relaxed and informal atmosphere is advantageous, and the serving of refreshments generates a more natural attitude on the part of the consumer and deflects her attention from trying to interpret what she is in the Laboratory to do. Experience also strongly suggests that attempts to make the various laboratory areas appear more "real" (for example, to design a "store" that really looks like a store) are certainly not damaging to the collection of meaningful data. But, more importantly, such attempts are only marginal in their benefit to the validity of the Laboratory Test Market system. Normal consumers know they are participating in some kind of an experiment, and no attempt to eliminate this knowledge can be effective. The "store" must be enough like a store for the consumer to be stimulated to adopt a "shopping" attitude; it must stimulate the kind of attention normally paid to the item on display in a store and to the purchase-decision process usually followed. (See the diagram of physical elements of the Laboratory Test Market in figure 9–1.)

Differences between the LTM Environment and the Real World

Although the store replica which is the LTM point of purchase is designed to encourage consumers to move into a normal "shopping mode," purchase behavior there normally differs markedly from what would be expected in the real world. Further, the period of in-home use (in the interval

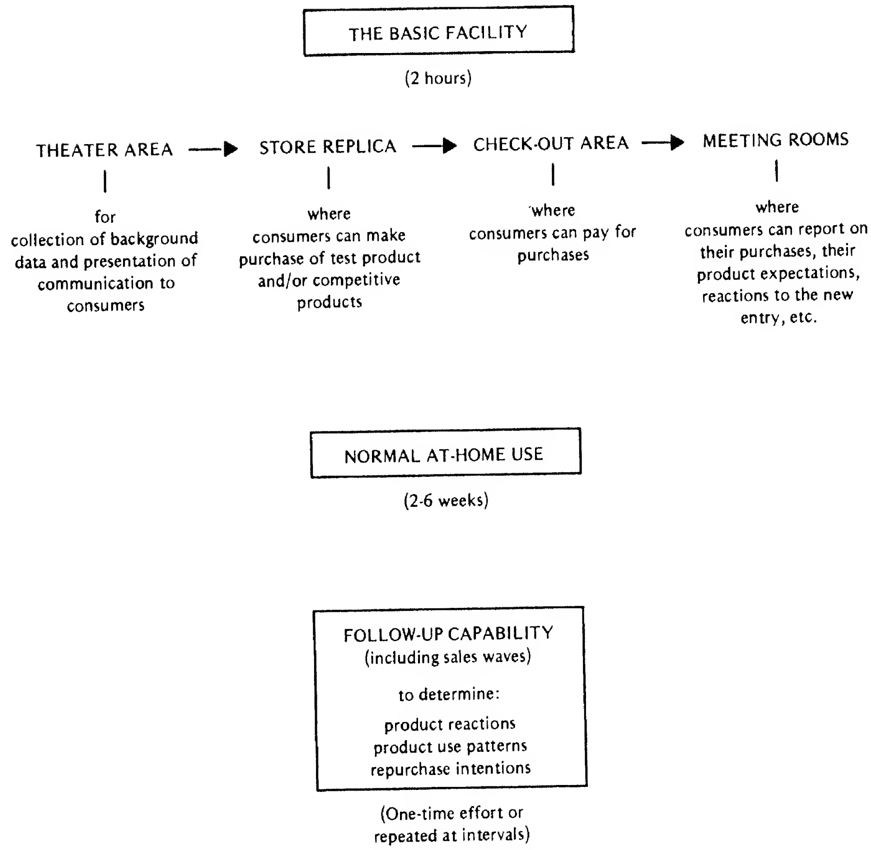


Figure 9–1. Physical Elements of the Laboratory Test Market

between purchase in the Laboratory store and re-interview) may also include some artificial effects which must be considered before undertaking the estimate. An understanding of how the LTM differs from the real world forms the basis for the LTM estimate.

Elevated Awareness in the LTM. As all marketers know, it is virtually impossible to attain 100 percent awareness of new products or of product modifications or line extensions. It often takes years for some products to become known to a majority of the public. By contrast, awareness levels in the LTM frequently approach 100 percent. This is true because LTM respondents comprise a captive audience in the sense that, when seated in the auditorium viewing the television vehicle with advertising inserted, they do not have the option of leaving the room during commercial breaks, as may

be their habit at home. Thus, while product advertising is realistically presented in the context of a television program, the condition of heightened attention which obtains in the Laboratory is equivalent to multiple advertising exposures in the real world.

Elevated Purchase Levels in the LTM. Heightened awareness due to the captive-audience situation, along with the reduction of time between exposure to the messages and the opportunity to buy (about 15 minutes to ½ hour in the Laboratory; as opposed to days and weeks in the real world), operates to significantly increase trial interest and subsequent purchase in the Laboratory store—far above real world levels. The degree to which laboratory trial rates are greater than in the real world varies, however, contingent upon a number of conditions. For example,

> Incremental trial rates due to the Laboratory situation are less marked for product areas which basically are not really interesting to consumers.
>
> Products with a "novelty" appeal—unusual, new, intriguing—even with no meaningful benefits—tend to generate extraordinarily high trial rates in the Laboratory Test Market, particularly if the cost is too low.
>
> Products whose marketing strategy depends on "humorous" advertising or on unique advertising execution techniques rather than on content showed less inflated trial rates in the Laboratory Test Market.

In addition, purchasing in the Laboratory is artificially stimulated. Since the Laboratory Test Market is a technique designed to compress the real world in *time* as well as in space—the purchase level data must reflect what consumers might normally do over several visits to a store, after awareness of a new product has been achieved. Hence, the system must strive to maximize purchase levels during the visit to the Laboratory store. Maximization—within meaningful parameters—is achieved through an incentive, usually money. The money—in coin—is given to the consumer as she enters the Laboratory store. The amount of money usually represents about 20 to 30 percent of the average price of the products available for purchase, i.e., the new product and its key competition. The coin, or "seed money," can be used toward a purchase—the balance to come out of the consumer's *own* money. If the consumer simply is not a user of the product category or is not influenced to make a purchase even by the seed money, she has the option of simply keeping the money.

Rate of Use among LTM Consumers. One of the strengths of the LTM design is that the products bought in the Laboratory store are taken home

by consumers and used in a normal at-home setting. Experience indicates, however, that the *rate* of use of new or modified products often differs from that found in the real world. More important for LTM purposes, consumers' expectations as to the *expected* rate of use of the test product are often unrealistic. This is particularly true for new products, especially those which may comprise a new product category. Projected use of products in existing categories, where a pattern of use has already been established, is generally reasonably accurate as reported by consumers. However, for entirely "new" products, confidence levels about future use patterns are lower.

Further, if the new product under study is of a type where wear-out can occur after an initial period of frequent or regular use, or of a type where usage/frequency patterns are established only after a period of experimentation, there is often the need for a mechanism where several opportunities for purchase and subsequent in-home usage can be provided. Accordingly, contingent on the product category, the Laboratory Test Market system includes research via sales waves. As is true of the Laboratory system in principle, the sales wave technique (no matter how many waves are involved and irrespective of the elapsed time for each wave) also represents a compression in time. This, too, is therefore regarded as an element of difference between the LTM and the real world.

The basic strategy of the LTM estimate is to remove the effect of these experimental effects from the LTM data through a set of corrective factors which serve to convert the LTM data to real world equivalents.

The LTM Estimate

Like the conventional test market, the LTM is designed to generate an estimate for a point in time when the sales rate has stabilized after the introductory period. This stable state may occur as early as six months after introduction, in the case of a frequently purchased product, or may take up to two years to achieve, in the case of products where the normal purchase cycle is relatively long.

The way in which this estimate is made from the LTM data described in the previous section is detailed in this section. More specifically, these subjects are covered: (1) an overview of the estimating process; (2) data inputs for estimating; (3) corrective factors applied to the data; and (4) estimating for the five basic LTM experimental models.

Overview of the Estimating Process

A rudimentary model of the basic LTM estimating process could be depicted in the following way:

LTM Purchase Rate

(% of all LTM respondents who buy test product)

×

LTM Repurchase Rate

(% of purchasers who will repurchase)

×

Frequency of Purchase

|

Estimate

This is a conceptual model. The working model must be adjusted to remove from the data the laboratory effect of elevated awareness/purchase levels cited in the preceding section, to "fine tune" frequency of purchase data, and to take into account the level of promotional effort planned to support the product at national introduction as well as anticipated extent of distribution.

The sections which follow explain the transition from a conceptual estimating model to a working one.

Data Inputs for Estimating

The first step in the estimating process is, of course, the collection of a body of specific data about the new product being studied. These data come primarily from the LTM research. However, some client input is needed, as well.

LTM Data. The Laboratory Test Market system collects data through at least four different methods, with a possible fifth method if sales waves are included in the study (dependent on the nature of the product). The different data collection methods are a self-administering background questionnaire, purchase records reflecting behavior in the laboratory store, quantitative and qualitative feedback following a visit to the laboratory store, call-back interviews after normal in-home experience with the product; and possibly the interviews conducted as part of sales waves.

Table 9–1 shows the specific information bits which must be collected

Table 9–1
Project-Specific Data Collected in the Laboratory Test Market Necessary to the Estimate

Type of Data	*Data-Collection Method*
Proportion of consumers who report using the product category Brand preference (prior to exposure to new product) Usage of product: frequency, regularity	Background questionnaire
Proportion of consumers who buy anything at all in LTM Purchase levels for new product Purchase levels for existing competitive products	Laboratory store records
Reasons for not buying any product in laboratory store	Focused group discussion
Reasons for buying new product	Postpurchase questionnaire
Reasons for not buying new product	Focused group discussion
Patterns of use of new product Frequency of use of new product Proportion who report commitment to continue to use new product Anticipated future usage patterns for new product (regularity, frequency) Reasons for satisfaction/dissatisfaction with new product	Call-back interview Sales-wave interviews

via each method in order for a Laboratory Test Market study to be successfully executed. (It should be noted that the need for these specific data governs the scope and design of the various questionnaires to be used in LTM studies and sometimes makes it difficult to add items to questionnaires which are of interest to marketers but which are not germane to the Laboratory Test Market system.)

Data Provided by Clients. In addition to the data inputs generated by the research, the estimating process requires some information from the marketer. These data are needed primarily because the LTM estimate must take into account the weight and character of the introductory effort and includes advertising weight, promotional activity, (e.g., sampling, couponing), and anticipated distribution level.

Corrective Factors Applied to the Data

To bring the LTM data into line with the real world, the estimating process requires that several adjustments be applied to the data in the form of corrective factors. Three such factors are used in LTM estimates:

Salience/Novelty factor: To remove the effects of elevated awareness and purchasing that are characteristic of the laboratory setting.

Clout factor: That portion of the estimating model which accounts for the level and type of introductory spending by the marketer, as well as the anticipated distribution of the new entry.

Frequency factor: To account for the frequency with which a product will be purchased.

The LTM estimating system is strictly an empirical one. The adjustment factors which are the core of the system are *not* derived theoretically. Each one comes as the result of repeated comparisons between LTM data and subsequently gathered real world data. Before the first LTM study was undertaken, these data comparisons had been in progress for some five years. Further, refinement of these factors is an ongoing process, and all LTM results are monitored with an eye to continual sharpening of the estimating system.

The Salience/Novelty Factor. That the observed purchase levels in the Laboratory would be inflated was known from the beginning as the concept of the Laboratory Test Market was being formulated. We knew, therefore, that a reduction factor to counteract this ''laboratory effect'' would be needed. The determination of just how much of a reduction was required in order to bring the LTM data into line was, of course, an empirical process begun during the LTM developmental period and refined in our early studies.

These studies led us to the early conclusion that LTM purchase data for test products were inflated by one-third, calling for a factor of 0.75 to be applied to our data, so that

$$\text{Real World Purchase Incidence} = \text{LTM Purchase Incidence} \times .75$$

Over the years that the LTM has been in operation, this relationship has held true for better than 9 of every 10 experiments. However, in the remaining 10 percent of cases, considerable variation has been noted. Although no instances have yet been noted where ''laboratory'' inflation was completely

absent—occasional cases have been encountered where LTM purchase levels were only about 10 percent above corresponding real world levels. At the other extreme, products have been tested where laboratory purchasing was inflated by as much as two-thirds.

The factor used in the LTM estimating process to account for ''laboratory'' inflation of purchase levels is termed the ''novelty/salience'' factor because it has been found to be a function of the novelty appeal of the product being tested and of the salience of the product category in which it resides. Products which represent a significant departure from the existing entries in a category, or which mark the start of a new category altogether, tend to generate high novelty appeal and are subject to highly inflated purchasing in the laboratory. Conversely, products which differ little from existing brands and those in low-salience categories which generally hold little interest for consumers are subject to relatively little exaggeration in the laboratory context.

As mentioned earlier, the empirically derived Salience/Novelty factor as applied in LTM estimates averages .75 (33% inflation), but ranges from a low of .60 (67% inflation) in high novelty/high category salience situations to a high of .90 (11% inflation) at the other extreme.

The Clout Factor. In several key respects, the Laboratory Test Market represents an optimal condition from the marketer's viewpoint in that brand awareness is nearly at the 100 percent level and because the purchase situation follows advertising exposure by only a few minutes. Further, the unfailing presence of the test product on the Laboratory store shelves replicates a condition of 100 percent distribution. By design, an out-of-stock situation cannot occur.

It is rarely true in the real world that a marketer's promotional clout at introduction is so great that a situation comparable to the LTM obtains. Accordingly, there is the need to further modify observed purchase rates in the laboratory. The adjustment used for this purpose in the LTM is called the ''clout'' factor because it is entirely dependent on the amount of promotional ''muscle'' which would support the new product were it to be introduced in the real world. Very heavy introductory spending accompanied by nearly full distribution would call for little ''clout'' reduction of observed Laboratory purchase levels. Conversely, an introduction of modest scope and spotty distribution would call for a rather drastic ''clout'' reduction.

It will be recalled that the LTM estimating process requires some data provided by clients (advertising weight, anticipated distribution level, etc.). This is necessary solely to arrive at the proper clout factor. In fact, it is at this stage of the LTM estimating process that there is frequently client involvement as the various components of the clout factor are jointly considered and discussed by LTM and client executives.

The specific considerations which affect determination of the "clout" factor are the following.

The absolute level of media expenditures: Generally, the greater the expenditure, the higher the "clout" factor (i.e., the less the reduction of observed laboratory purchase levels).

Relative level of media expenditure: To a degree, competitive media expenditures represent "static" above which the client's voice must be heard. Therefore, high media expenditures relative to competition tend to raise the "clout" factor while low level spending serves to reduce it.

Absolute and relative levels of promotional expenditure: As is true for media spending, high promotional expenditures work to increase the "clout" factor, while lighter spending acts to reduce it.

Character of promotion. Some types of promotion have proved to be highly effective in generating awareness and stimulating initial purchase in certain product categories. For example, sampling has been widely effective in introducing personal and laundry products. Similarly, cross-couponing has worked efficiently where a good fit exists between the new product and a strong, established product. Thus, the choice of promotional technique for the particular category involved is also a factor in determining "clout."

Distribution: Where the test product fits into an existing category, the depth of distribution relative to established entries is a key "clout" component. To the extent the new product fails to achieve the average level for the category, it will be penalized by a lowered clout. On the other hand, the "clout" factor will be raised if anticipated distribution exceeds the category average. In the case of a unique new product which represents the start of a new category, the absolute level of anticipated distribution is considered.

Brand visibility/accessibility: There are often instances where new products find themselves at a distinct disadvantage in the supermarket, despite having achieved distribution, because their visibility or accessibility is limited. For example, a semi-moist dog food bulk packaged like a dry food and displayed in the dry dog food section is at a disadvantage because it can be (and was) mistaken for a high priced dry dog food. And there are, of course, instances where new products have been overshadowed by big selling brands with far more facings. Such visibility/accessibility considerations also come into play in determining "clout" factors.

Maturity of the market: A new entry in a mature market, one which has been stable for a number of years with little relative change among competing brands, is harder to "crack" than a growing market where successful new entries have emerged. Thus, the maturity of the new market becomes an issue in determining "clout."

Market dominance: Categories dominated by a single brand, where other entries are only minor factors, tend to be more difficult to "crack" for a new product. Here, purchase habits are well set and almost habitual. In such categories, more clout is needed to achieve the same result as can be

obtained with less in a volatile market where multibrand usage or switching is more common.

Experience in the category: A marketer entering a category which is unfamiliar to him is likely to make mistakes that would not be made by the marketer who has been dealing in the category for years. Thus, experience is also an element in determination of the "clout" factor.

To arrive at the final "clout" factor, all these elements are considered carefully, often in conjunction with client personnel, and evaluated in the light of our large body of normative data developed over the years of Laboratory Test Market operation. The "clout" factor represents the degree to which observed LTM purchase levels must be reduced because of the scope of the marketer's introductory efforts. Generally speaking, a heavy effort calls for a small reduction while a light effort leads to a substantial one.

In practice, "clout" factors range from .75, a 25 percent reduction of LTM purchase rates, to .25, a 75 percent reduction appropriate to a light introductory effort. These specific percentage corrections are not theoretical, but empirically derived through client feedback and subsequent analysis. However, it must be noted that, in actual fact, neither the Laboratory Test Market system *nor the real world* can validly measure the effect of small (in terms of consumer impact) differences in introductory weight. For example, our data resource suggests that the corrective factor indicated by the following two plans is likely to be the same:

Plan A:	Advertising level	$2,500,000
	Distribution depth	60%
Plan B:	Advertising level	$3,000,000
	Distribution depth	65%

The fact that these two plans will yield the same corrective factor when the LTM data resource is consulted is not *theoretical;* it is a function of empirical feedback on the thresholds needed to achieve a differential impact in the marketplace.

Frequency Factor. The application of the Salience/Novelty and "clout" factors permits an unbiased estimate of the incidence of use for products tested in the Laboratory Test Market. Often, no additional factor is needed. In some instances, however, the application of a frequency factor is crucial.

The need for such a factor was not immediately apparent during the early years of the Laboratory Test Market. The first 25 LTM validations suggested that our estimating system was remarkably accurate. However, the 26th estimate proved to be high. This led to a period of reevaluation of the system and to the subsequent realization that, while estimates for new

products in existing categories were on target, those for new products which, in effect, began new categories were more difficult.

In an established category, a new entry acts largely to displace other brands. In such cases, frequency of use is not a problem since a pattern of use for the category is not difficult to establish and any need for minor frequency adjustments for the new entry is usually apparent from the data. In other words, it is relatively simple to determine whether the rate of use of the new product will be on par with the category average or whether it will be somewhat high or low.

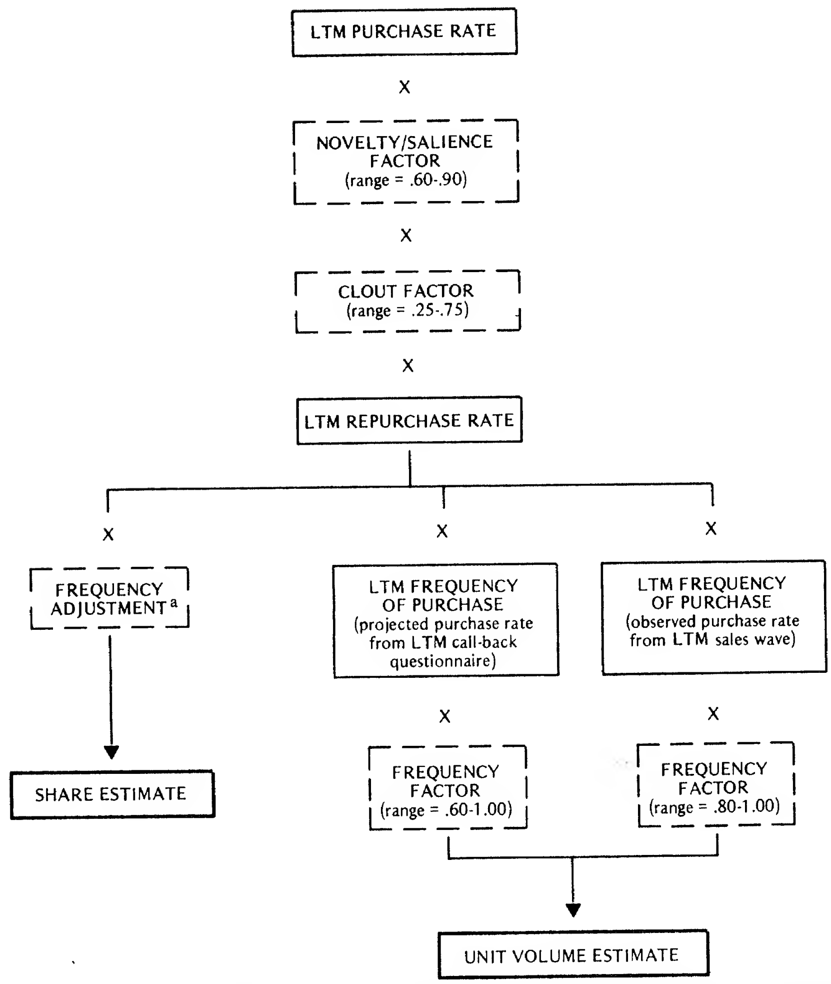
In a new category, where a pattern of use has not been established, there are no reliable benchmarks upon which frequency estimates can be built. Thus, in our early LTM work, the frequency of use component in LTM estimates was based on consumers' reports as to how often the product they had purchased in our Laboratory store would be used if it were freely available. Recognizing that such reports would be subject to exaggeration, an allowance for overstatement was made. However, exaggeration was greater than anticipated, leading to our overestimate in the 26th LTM validation.

Since that time, our ability to deal with exaggeration in consumers' reports of projected frequency of use has improved markedly. In addition, we developed a sales wave technique whereby reorder of the test product is periodically solicited from LTM respondents, so that a behavioral measure of frequency of use became available.

Continual refinement of both the consumer reporting and sales waves techniques for establishing frequency of use has brought us to the point where alternative frequency factors for both methods are available in LTM estimates for new products which do not benefit from a category frame of reference. As with all of our estimating factors, the frequency factors are empirically derived from comparisons of LTM data with subsequent real world data. In current studies, frequency factors are selected on the basis of our normative data. Where sales waves are used, the frequency factor ranges from 1.00 (no reduction of observed frequency) to .80 (20% reduction). Where frequency is based on consumers' reported intentions, the factor ranges from 1.00 (no reduction) to .60 (40% reduction).

Application of Factors in Estimates

Laboratory Test Market studies are of five distinct types, each with a different estimating model (described in the following section). The general model shown in figure 9–2 is offered simply to show how the corrective factors are applied in the estimating process. The figure shows two alternate end-products of the estimate—unit volume and market share.



[a]Applied only if data indicate test product will be used at a rate different from category average.

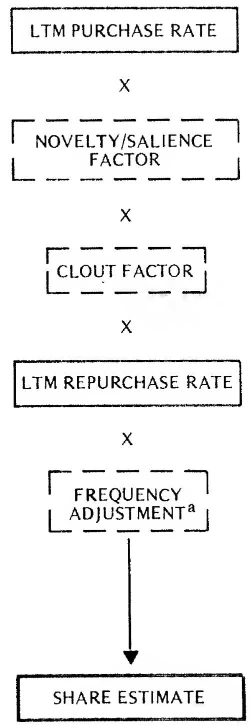
Figure 9–2. General LTM Estimating Model

Five Estimating Models

Laboratory Test Market studies are of five distinct types with the choice of types a function of the nature of the test product, the nature of the product category, and the specific needs of the client. Each of these types employs a different variation of the general estimating model shown in figure 9–1. The tables and figures which follow present a brief description of each study type and the estimating model applicable to each.

Table 9-2
Standard LTM

Application:	New brands which are positioned as replacements for existing entries in established categories
Estimate:	Generally, share of category units, dollars or users



[a]Applied only if data indicate test product will be used at a rate different from category average.

Figure 9-3. Standard LTM Estimating Model

Table 9-3
Extended LTM

Application:	New products which are additions, rather than replacements of existing entries (for example, candy products)
	New products which do not fit into existing categories and, in effect, comprise a new category (for example, new foods)
	New products which may be subject to wear-out (for example, snack foods)
Features:	Sales waves used over an extended period to obtain behavioral data on frequency of use
	Permits observation of trend in frequency of use over time
Estimate:	Usually unit volume

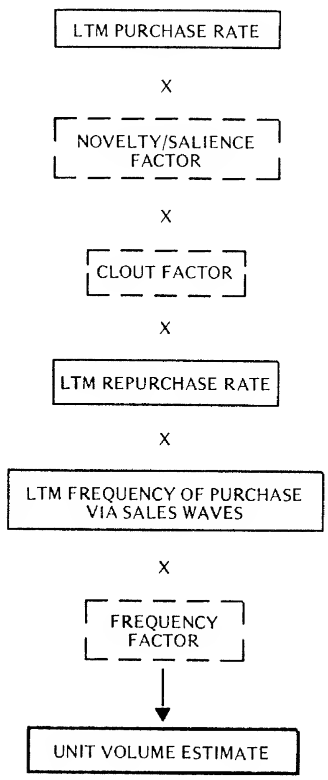
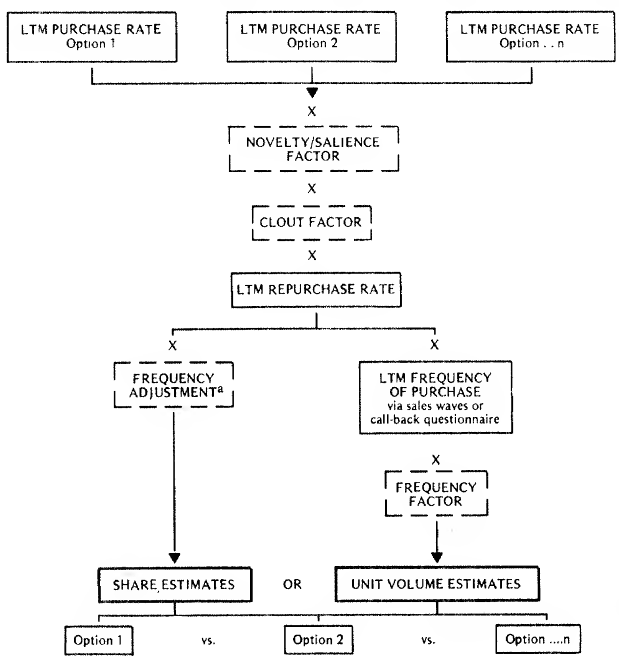


Figure 9-4. Extended LTM Estimating Model

Table 9-4
Strategic LTM

Application	New products where two or more strategic options are being considered; options may be in alternative Pricing Advertising Store position Packaging Etc.
Features:	Multicell experiments Sales waves, if appropriate
Estimate:	Independent estimates for each experimental cell May be in terms of unit volume, or share of users, dollars or units

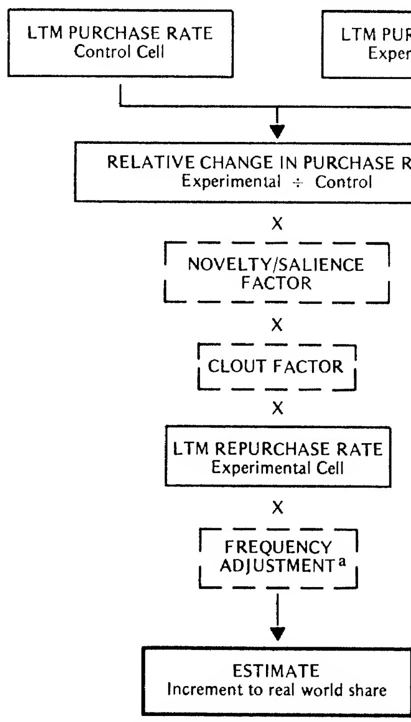


[a]Applied only if data indicate test product will be used at a rate different from category average.

Figure 9-5. Strategic LTM Estimating Model

Table 9–5
Modification Laboratory Experiment (MLE)

Application:	Modification of existing product, such as line extension; packaging; dispensing mechanism; price change
Features:	Multicell experiment including a control cell
	Analysis in line-extension studies focuses on user analysis to determine if "cannibalization" will occur
Estimate:	Increment to real-world share

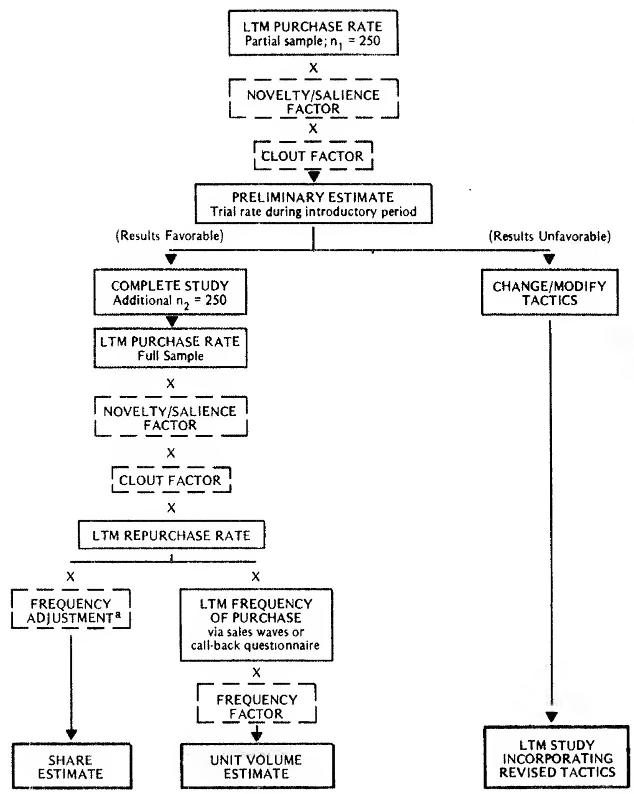


[a]Applied only if data indicate significantly different usage rate in control and experimental cells.

Figure 9–6. MLE Estimating Model

Table 9–6
"FACTS" LTM (Fast Assessment: Chance to Temper, Transform Strategy)

Application	New product in relatively early stage of development
	Marketing elements not yet firmly in place
Features:	Early estimate of trial based on half sample
	Early estimate the basis for determining whether modification of strategy/product is required or whether experiment should be completed
Estimate:	Preliminary estimate (based on half sample) of trial rate during introductory period
	Full estimate (if study carried on to completion) in terms of unit volume, or share of users, dollars or units



[a]Applied only if data indicate significantly different usage rate in control and experimental cells.

Figure 9–7. FACTS LTM Estimating Model

10 COMP: A Comprehensive System for the Evaluation of New Products

Philip C. Burger, Howard Gundee, and *Robert Lavidge*

Overview

Anticipating consumer reactions to a new product after it has been introduced in the marketplace is vital in developing effective and efficient marketing plans. In particular, forecasting the volume of consumer purchases is critical in deciding whether to market the new product and, if so, in producing and distributing optimum quantities of the product for its introduction.

To minimize the risks involved in introducing a product on a broader scale, test markets often are employed. Experience has demonstrated that they can be very useful. However, it also has been demonstrated that they frequently have significant limitations. (1) A period of six months or more commonly is needed (often, the required time is more than one year). (2) Costs are high. (3) Competitors frequently take advantage of test marketing to plan the competitive actions.

The new product evaluation system described in this paper discusses an alternative means of evaluating new products that overcomes or minimizes these test market limitations. The system, called COMP, was developed in 1971. Since then, approximately 80 tests have been completed. In some of these cases, the products evaluated have not been introduced into the marketplace. In other instances, the experience in the marketplace has not yet been sufficient to determine whether the COMP sales predictions were valid or invalid. However, in each of the more than 20 instances where it has been possible to determine the accuracy of the COMP predictions, those predictions have proved to be remarkably precise. Moreover, the diagnostic information developed through the comprehensive evaluation process has proved to be of value in improving the marketing plans prior to the introduction of some of the products evaluated.

The COMP system, as such, is designed to evaluate consumer products of the type customarily bought on a self-service basis. It does not provide sales predictions for products which require personal sales effort. However, parts of the COMP system have been used successfully to diagnose the sales appeals and purchase deterrents of such products and services.

The Model

The theory of consumer behavior central to the COMP system is based on a variation of the model proposed by Lavidge and Steiner (1961) and results reported by Pessemier, Burger, and Tigert (1967). For a product of the type for which the COMP system is considered appropriate, the probability that a potential purchaser will buy it is assumed to be related to awareness of the product's existence and the individual's perceptions of it. The model uses measures of each buyer's attitudes, purchase behaviors, and repurchase intent and relates those measures to the anticipated introductory advertising and distribution program to predict a probability of purchase. The predictions for individual consumers are aggregated to predict total market share.

The underlying hypothesis of the model is that consumers develop a set of attitudes and intentions toward a given subset of products in a particular market. Consumers are viewed as having one "most used" brand for which there are well-defined attitudes. In addition, consumers have evaluative information about the major brands in a market. All brands are evaluated in terms of the evaluative criteria compared to the most used and the largest-share (benchmark) brands. Thus, new brands are compared to these brands. If the new brand is evaluated as being preferred to the most used and benchmark brands for a great number of consumers, the new brand should achieve substantial market share if it is marketed properly.

The model can simulate the impact of a variety of management strategies. First, it can, by appropriate selection of data, predict the impact of a sampling campaign upon consumer acceptance. Also, the model predicts market share for a range of distribution and awareness levels. Price and packaging variations can be tested by using multiple-cell (panel) designs. The potential effect of revisions in product formulations can be evaluated by computing the market shares that would result if various changes in consumers' perceptions took place. Thus, the system can provide substantial relevant information to the marketing manager.

Experimental Procedure

Selection of Participants

The sample of consumers selected to participate in the experimental procedure must include users and potential users of the type of product being evaluated. However, the essential product usage and demographic characteristics of persons to be included may be defined on the basis of either specified target markets (those at whom the product is being aimed) or a

broader universe, to aid in determining which groups are the best prospects. (This procedure is most likely to be used when the product does not fit into a well-defined product category.)

Screening interviews then are conducted, normally in shopping centers, to recruit an adequate number of the appropriate kinds of consumers. They are selected one at a time, never from groups, to avoid participant interaction. The consumers are paid a nominal sum (usually $1.50 each) for their participation in the experiment.

Measurements prior to Exposure to the New Product

The selected consumers are interviewed prior to exposure to the new product, or advertising for it, to obtain measurements of such factors as awareness of existing products or brands, relevant purchase and use behavior, future purchase intentions, additional demographic factors, product attribute importance ratings, and perceptions (attribute ratings) of the products included in the experiment as controls. Respondents always rate the brand they most often use and those benchmark brands with which they are familiar.

As part of the initial interview, each respondent is questioned to determine awareness of benchmark brands. Subsequently, the respondent is asked to rate a set of benchmark brands in addition to the test product. The benchmark brands always include the brand which the individual names as the one used most. The most used and benchmark brands are rated on a series of benefit and attribute scales prior to exposure to the experimental design.

Advertising Exposure

After the initial data are gathered, respondents are exposed to television or print advertising for the new brand and the benchmark brands. If both television and print advertising are used in marketing the benchmark brands, or if they are to be used in the introduction of the test product, the respondents may be exposed to both media. Normally, this exposure is limited to the television commercials or print advertising. The objective is not to determine the strength of the advertisements in attracting attention. Rather, it is to transmit the message as set forth in the advertising. To reinforce the exposure, each respondent is questioned about the advertising. This questioning also provides useful information about the impact of the various elements included in the commercials and/or advertisements.

Initial Exposure to Packaging and Pricing

Following exposure to the advertising messages and questioning about them, each participant is given an opportunity to purchase the experimental product and/or any of the control items at low but realistic prices in a simulated retail store section.

Shortly before the beginning of a COMP study, stores near each Test Center are visited to determine the prices at which the benchmark brands are being offered. Those brands usually are made available in the Test Center at the lowest prices reported, excluding "specials." The prices at which the test brands are offered in the Test Center are established in relation to these benchmark prices. To evaluate the impact of alternative prices at which the test product might be offered, additional panels of test participants who are exposed to them may be used.

When the purchases, if any, have been made, each participant who does *not* purchase the new brand is given a free sample of it to increase the efficiency of the use-test phase of the experiment. To minimize bias which otherwise might result, each consumer who buys the new brand is given a free sample of an unrelated item of approximately the same value. Each respondent also is given a silver dollar and a note thanking her/him in advance for continued cooperation in the test and informing her/him of the subsequent interview to be made.

After-Shopping Interview

Before respondents' leaving the Test Center, one additional questionnaire is completed by the test participants. This interview is similar to the questionnaire completed prior to the advertising exposure. Its purpose is to secure brand ratings and purchase intentions for those brands not previously rated and for the test brand. Respondents at this point in the procedure are able to react to these brands based on the advertising, packaging and pricing to which they have been exposed.

After-Use Measurements

When adequate time has passed for the participants to have tried the products secured in the simulated retail store, they are recontacted in person in their homes. Other sales prediction systems are known to use dissimilar data collection techniques—for example, using self-administered forms in the Test Center to collect rating data and later obtaining this same type of data via telephone after-use. Personal, in-home interviews are conducted (as

opposed to telephone interviews) in order to maintain comparability with the data collected in the Test Center.

The in-home interviews provide information about reactions to the test and control products after the purchased products have been used. Data collected includes future pruchase intentions, satisfactions and dissatisfactions with the products they have purchased or received as samples and perceptions of the test and control brands included in the experiment.

The Data Analysis and Prediction Model

A number of models are used in deriving the sales prediction. The models include (1) the attitude paradigm which combines importance ratings and brand ratings into a single score for each brand in the study, (2) the purchase intention transgeneration, (3) the regression model which expresses the relationship between purchase likelihood and attitudes and purchase intentions and (4) the market share prediction.

The market share prediction model uses attitude and purchase intention data collected after exposure to the new brand plus regression results to compute the impact of the new brand. Specifically, the market share model contains four distinct options:

1. Market share is predicted under conditions simulating a marketing strategy where advertising is the major portion of the promotional mix.
2. Market share is predicted under conditions simulating a marketing strategy which combines advertising with trial stimulating programs such as couponing or sampling.
3. Market share is predicted in either of the above ways assuming static market conditions. (Market share for the new product will be at the expense of existing products.)
4. Market share is predicted in either of the above ways but providing for the possibility that the new product will attract non-users of existing brands by overcoming major objections of nonusers. Thus, the new product achieves sales not only at the expense of existing brands but also as a result of non-users becoming users.

Preference-Belief Submodel

A key proposition in the consumer model relates the hypotheses that for any given person, probability of purchase of a given product or brand is influenced by the relationship of the product or brand's rated attributes to the attributes of the product or brand bought most often. Specifically, if a new

brand is rated as being similar to the most used brand, it is likely that the purchase probability will be proportionately close to that of the most used brand. In addition, each respondent weights his or her attributes by importance. Thus, if there are 15 attributes, each attribute is weighted by each respondent in terms of its importance in making purchase decisions within the product category (Wilkie and Pessemier 1973). This importance rating is designed to approximate the utility profile of each attribute. The attributes are defined to be specific, well-understood product benefits and characteristics.

Mathematically, the model is constructed in the form

$$A_{ij} = \frac{\sum_k I_{ik}(MU_{ik} - B_{ijk})^2}{\sum_j \sum_k I_{ik}(MU_{ik} - B_{ijk})^2}$$

where I_{ik} is the attribute k's importance rating for person i, MU_{ik} is the rating of the most used brand on attribute k, and B_{ijk} is the rating of brand j on attribute k for person i. This particular form has been suggested by the literature of multidimensional scaling (Green and Carmone 1970). The numerator is a computed score multiplying the squared score of attribute ratings of the most used brand minus that of brand times the importance of the attribute. The results of each attribute score are summed over all attributes. The denominator is the score summed across all brands rated. This summed score serves to remove scale bias (Bass and Wilkie 1973).

Likelihood of Purchase Ratio

The behavioral measure used for purchase is the ratio of reported purchase of brand j within a defined recent past period. The questioning varies from product category to category, but the data are cast into the form

$$P_{ij} = \frac{\text{brand}_{ij}}{\sum_j \text{brand}_{ij}}$$

where brand_{ij} is the proportion or frequency of purchase of brand j by person i, and P_{ij} has range 0 to 1 and metric properties. It is used as a surrogate for purchase probability.

The Purchase Intentions Submodel

Purchase intentions are collected at several points in the COMP experiment. These intentions are 5-point semantic differential scales where the number 0

represents "definitely would not buy" and the number 4 "definitely would buy." Since semantic differential scales are subject to spurious scale variance, the scale scores for each brand under study are normalized by the following manipulation:

$$IN_{IJ} = \frac{IN_{ij}}{\sum_{j} IN_{ij}}$$

Regression Model

Ordinary Least Squares (OLS) regression is performed to test the attitude-purchase probability relationship. The analysis is performed pooling across all subjects in the sample who report valid data (missing subjects are deleted). The regression results for each brand are obtained independently. Thus, there is a separate regression analysis across all subjects for each brand. The basic linear hypothesis is

$$\ln P_{ij} = a_j + b_{1_j} \ln A_{ij} + b_{2_j} \ln IN_{ij} + e$$

In performing log-linear regression with bounded dependent and independent variables, each end of the curve is "anchored." The major problem becomes estimating where the "elbow" in the curve is. The key data needed to estimate the elbow are provided by those subjects whose data occur around the elbow. Behaviorally, the location of the elbow helps to understand how similar a brand has to be to the most used brand before it moves from the nonpurchased category to the purchased category. The results from performing the regressions are shown in table 10–1. Quite different product categories are reported. Analgesics are a category where there is low multiple brand purchasing. Branded cosmetics are quite different. There is a great deal of cross-brand purchasing in cosmetics because of the wide variety of colors and fragrances offered within a product line. Thus, the cosmetic raw data included a large number of fractional purchase probabilities. The results show quite good fits in both categories. The R^2 and F's are extremely significant. In addition, the b_j coefficients have the proper sign (negative implies that the farther from the most used brand, the less the probability of purchase). Data for a branded beverage also are shown for comparison.

The average regression results for the analgesic brands were plotted in figure 10–1 to show the shape of the curve. The purchase probability drops radically as A_{ij} moves toward 0.2. Past the 0.2 level, there is almost no chance that a person will buy brand j. In terms of the raw attribute ratings, it can be deduced that relatively small rating differences on important attributes can substantially reduce probability of purchase.

Table 10-1
Regression Results

Brand	b_j	R^2	F[a]
	Analgesic Regression Results		
1	−0.81	.79	188
2	−0.41	.32	23
3	−0.52	.42	356
4	−0.36	.31	225
	Branded Cosmetic Regression Results		
1	−0.81	.54	586
2	−0.47	.22	142
3	−0.33	.41	342
	Branded Beverage Regression Results		
1	−0.91	.46	16
2	−0.93	.54	23
3	−0.08	.05	100

[a]All significant at p less than .001.

The New Brand's Regression Values

Clearly, there is no historical purchase information for the new brand under study. Most often, to compute the regression parameters for the new brand which will be used for forecasting purposes, the average values of each a_j and b_{ij} are computed using the estimates for all the benchmark brands. This assumes that the new brand's attitude, intention, and purchase relationship will be similar to the average of the benchmark brands in the category.

The Market Share Simulation Submodel

The underlying hypothesis regarding market share is probabilistic in nature and based on the projected probability of purchases of individual consumers. The major hypotheses include:

1. Market share depends on both individual probability of purchase and usage rates, e.g., it is possible that high usage individuals may prefer the new product while low usage individuals do not. Thus, the product may be successful and yet have low average preference. The reverse also can be the case.

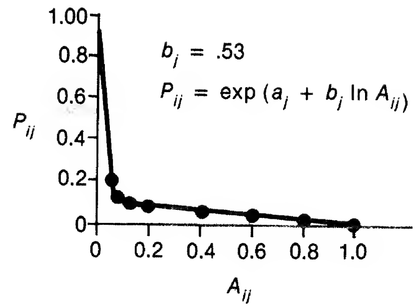


Figure 10-1. Response of Purchasing to Changes in Distance from Most Used

2. Probability of purchase depends on the level of actual distribution achieved after introduction. The weighted market share is reduced by the percentage All Commodity Volume (ACV) distribution expected to be achieved by management.
3. Probability of purchase depends on unaided but probed brand awareness. Unaided, probed awareness is defined as follows:

   Brands used most often
   Other brands used during a relevant time period
   Other brands used in the past
   Brands which would not be used
   Other brands of which respondents have seen or heard

   The market share is reduced by a function of the percentage of the population of consumers who are aware.
4. The mathematical form of the basic market share model can be written as

$$MS_j = \frac{\sum_i P_{ij} U_i}{\sum_i U_i}$$

Where MS_j is the market share of brand j, P_{ij} is the probability of purchase of person i for brand j, and U_i is the usage rate of person i.

The distribution function is linear; e.g., 40 percent is 0.4. Awareness was found to be a modified logistical curve. The parameters of the curve were estimated based on the results of brand usage and unaided awareness of 79 existing brands of frequently purchased products ranging from peanut butter to deodorants. The resulting curve asymptotes at 0 on the low end and 1.0 on the high end with a shallow S shape.

Since it is unlikely that the actual market share of the sample will exactly estimate national market share, a mechanism for correction has been developed within the system. This correction mechanism adjusts the intercepts of the individual regression equations

$$\ln P_{ij} = a_j + b_j \ln A_{ij}$$

in the following form:

$$\min_{a_j} [\sum_j (MS_j - \widetilde{MS}_j)]$$

$$\text{subject to } j = \text{not new brand}$$

where $\widetilde{MS}_j$ are the actual market shares of existing brands and where MS_j is the computed market share.

This non-linear minimization problem is solved by a direct search algorithm. Usually convergence occurs within two iterations.

Trial-Repeat Estimation Submodels

The model of the trial of the new product is based on the hypothesis that a person is likely to be a trier of the new brand if he or she buys it when given the opportunity in the simulated store *or* if the person develops perceptions about the new product of equivalent strength to her/his perceptions of her/his most used brand after seeing the advertising, pricing, and packaging for the new product. Thus,

$$T = \frac{B_a + E_{bb}}{N}$$

where T = trial percentage

B_a = actual number of buyers of the new brand in simulated store

E_{bb} = number of nonbuyers with equivalent preferences (of buyers) before actual new brand use

N = total sample size

This particular formulation is based on two considerations: (1) those who buy in the lab are showing behavioral evidence of curiosity and/or favor for the new product and (2) those with equivalent attitudes are likely to purchase but, for home inventory or other reasons, chose not to do so at the time.

The logic of the repeat purchase model is similar. It is hypothesized that, after using the new product, if the new product achieves equivalent or stronger perceptions than the most used brand, the individual probably will be a repeat purchaser. Mathematically, this is defined to be

$$R = \frac{E_{ba}}{t}$$

where R = repeat percentage

E_{ba} = number of respondents with equivalent strength perceptions after use

t = number of triers

Analysis of Purchase Intentions—A Validity Check

Each test participant is asked to indicate the probability that she/he will purchase the test product after having an opportunity to see the advertising for it and being exposed to its pricing and packaging. A similar measure is secured at a later date after the individual has had an opportunity to try the product. In addition to being evaluated together with similar purchase intention measures with respect to the benchmark products, the purchase intentions measurements for the test product are analyzed separately. Based on experience with similar measurements related to consumer products varying from low-price non-durable goods sold on a self-service basis (for example, salad dressing or shampoo) to relatively high-price durables (for example, color television sets), weights are introduced to derive additional estimates, based solely on the purchase intentions data, of the probability that each individual will purchase the test product. At one extreme, weights of 0.8 for "top box" ratings plus 0.3 for "second-box" ratings have been found to be appropriate when analyzing data related to low-price, frequently purchased, low-perceived-risk products which have never been tried by the individual. For more expensive, less frequently purchased, higher-perceived-risk products, the weighting factors developed through other consumer studies are as low as 0.28 for "top box" scores.

In addition to being used to provide a check on the trial purchase predictions derived from the preference model, similar weights are used to estimate the probability that each of the individuals likely to make a trial purchase will become a repeat purchaser after using the test product. This calculation takes into account the interaction between each individual's purchase intentions ratings prior to use and following use of the product.

Brand Switching

An important aspect of the sales prediction result is a brand switching analysis. This can be critical if there is a possibility of cannibalism from other products of the same company which commonly is the case when the new product is an extension of an existing product line.

Diagnostics

A major aspect of the analysis resulting from this comprehensive system is the diagnostic evaluation of the competitive strengths and weaknesses of the test product. This comes from an analysis of the attitude scale data. It is possible to evaluate the strengths and weaknesses of the test product and also to position it in relation to benchmark brands. Analysis of the rating scales also may yield information about the impact on sales of specific product qualities and attributes.

Market segmentation analysis may be performed by using the attitude scales. A factor analysis is used to identify the key attitudes and the number of respondents in the sample exhibiting these attitudes. This procedure is helpful in isolating the key attitudes which best distinguish between various user groups.

Prediction Results

Figure 10–2 shows a comparison of the results predicted in a number of COMP studies versus those actually achieved after the products subsequently were introduced. As shown, across the range of validated predictions, COMP has proved to be remarkably accurate, exhibiting no tendency to either over-predict or under-predict at the various share levels.

Table 10–2 illustrates the relationship between the proportions of people purchasing in the Test Center store and the predicted and subsequently validated market shares. The table shows that the examination of only store purchase data may be misleading. In a number of cases, validated market share predictions have exceeded by a substantial margin the proportions of people observed buying the test product in the Test Center. Conversely, in other validated studies, high observed purchasing levels for the test product have not resulted in high market shares.

Of course, many factors can affect the level of purchases made in the Test Center store. For example, products which are frozen, refrigerated or bulky likely will have lower purchase levels than smaller, shelf stable products. Moreover, purchase of some products is influenced to a great extent by seasonality. Examples include the purchase of such products as cold remedies and laxatives.

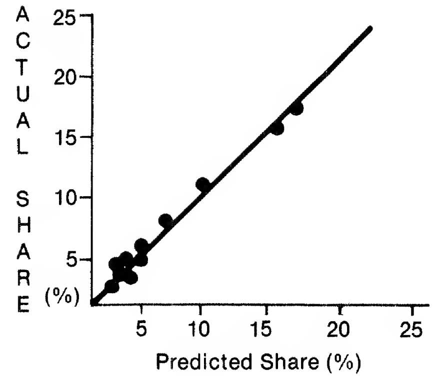


Figure 10–2. Predicted Share versus Actual Share

In addition to the results previously discussed, the COMP system provides a significant amount of useful diagnostic information, including such things as characteristics of triers, characteristics of repeat purchasers, product appeals, purchase deterrents, ideas communicated by the advertising, product use behavior, and effect of trial stimulation promotions.

Conclusions

This research design uses management science models and techniques to predict sales for a new product. It incorporates a simulated retail sales

Table 10–2
Relationship between Sales in the Test Center and Market Share
(percent)

	Purchased in Test Center	*Predicted Market Share*	*Validated Market Share*
Foods and Beverages			
Product 1	40	4	5
Product 2	5	17	17
Personal-Care Products			
Product 1	9	11	11
Product 2	29	5	5
Proprietary Drugs			
Product 1	32	7	7
Product 2	28	4	4
Household-Cleaning Products			
Product 1	40	16	16
Product 2	17	3	3

experiment and "in-home" use of the new product and provides a detailed evaluation of the product concept and use performance. Useful diagnostics are provided through the examination of the attitude scales. The information yield is rich in comparison to the cost of execution. Moreover, the likelihood of competitive disclosure is much less than with actual retail sales experiments.

Dependable forecasts of a new product's sales are of immense value to marketing management. However, they are not enough. In addition to knowing how much of a product will be purchased if the proposed marketing plan is followed, information is needed to aid in understanding the factors which are leading to its success or failure. This information can be of value in modifying marketing plans to improve on the projected sales volume. Equally important, it can play a vital role in preventing decisions to "kill" new products which could be sold on a profitable basis if appropriate changes were made in the manner in which they were marketed.

The use of this research design has been limited primarily to experiments involving products of the type commonly sold through supermarkets or drug stores which are dependent for their success on repeat purchasing as well as their initial appeal. To date, there have been no known estimation failures, i.e., no instance in which the sales estimate derived through the system differed substantially from reliable sales data available through other sources (such as Nielsen audits or factory shipments adjusted for wholesale and retail inventories).

We believe that combining the behavioral data derived through the simulated store experiment with quantifiable data relating to attitudes toward the experimental product and competitive alternatives has a synergistic effect. It increases the power of the sales forecast. In addition, it provides a significant amount of information of value in understanding purchase motivations and deterrents and, hence, in developing more effective and more efficient marketing plans for the new product.

References

Bass, Frank M., and William L. Wilkie (1973), "A Comparative Analysis of Attitudinal Predictions of Brand," *Journal of Marketing Research* 10 (August):262–369.

Green, Paul E., and Frank J. Carmone (1970), *Multidimensional Scaling and Related Techniques in Marketing Analysis,* Allyn and Bacon, Inc., Boston, Mass., 26.

Lavidge, R.J., and G. Steiner (1961), "A Model for Predictive Measurements of Advertising Effectiveness," *Journal of Marketing* 25 (October):59–62.

Pessemier, E.A., Philip C. Burger, and Douglas J. Tigert (1967), "Can New Product Buyers Be Identified?" *Journal of Marketing Research* 4 (November):349.

Wilkie, William L., and Edgar A. Pessemier (1973), "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Resarch* 10 (November):428–441.

11 Pre-Test-Market Research of New Packaged-Goods Products—A User Orientation

Joel Levine

Probably the least understood tool in the marketing-research field is pre-test-market research on new products. Over the years, in spite of serious technical issues, concept testing, product testing, advertising research, and even test marketing have developed an acceptance among management that in some cases may not be totally justified. Pre-test-market research is still emerging as a tool of the marketing-research profession that management can use (and many do) but the methods have not yet been fully validated.

The purpose of this brief chapter is to describe the uses that pre-test-marketing research can be put to, outline some of its limitations, and briefly describe the use of the Hendry model as an alternative approach to simulation test-market procedures. (For a more detailed discussion of the use of the Hendry model for forecasting new-product performance, see chapter 21 of this book.)

Use of Pre-Test-Market Research

It has been said by many (and I agree most emphatically) that test market should be considered "the last step in product development rather than the first step in the national introduction of a new product." That is, a test market is a structured, well-designed experiment to measure the impact of one or more national marketing programs for a new product. Key to the development of a marketing program is a national-volume expectation for the new product. Without a volumetric base to a marketing program, it is difficult, if not impossible, to develop a national spending plan for testing. It is certainly true that many develop a national marketing plan for testing in a test market by guessing how much advertising and sales-promotion dollars are required for introduction and then using some standard or corporate-wide goals to project backward to a volume estimate. Frequently, little is done at that point to evaluate the probability of attaining this volume objective. In my experience, it is desirable to have an independent estimate of the volume the new product can be expected to generate on the national basis,

based on explicit assumptions about the effects of marketing spending levels. A financial analysis (return on investment and so on) should then be conducted to determine if the expenditure level is acceptable. Key is the ability to develop accurate pre-test-market estimates of volume.

The objectives, therefore, of a pre-test-market research program can be classified into four areas:

1. To develop volume estimates for test-market planning. If the volume estimates are too low to support the business proposition, the entire program should be recycled. Further, it has happened that the pre-test-market prediction produced volume estimates substantially larger than the ongoing business proposition. This indicates the possibility of a larger business than originally thought and probably would result in planning additional test markets to estimate the upper bound of the proposition.
2. To provide a disaster check, since this is the first time a total marketing plan is exposed to the consumer. Packaging, advertising, pricing, and shelf locations are generally part of a pre-test-market program, and these elements of the marketing mix can be evaluated.
3. To provide information for modification of marketing programs prior to test marketing. Pricing levels can be tested as well as brand strategy, advertising strategies, and shelf location.
4. To obtain diagnostic information on product performance, the impact of advertising, and pricing.

To the extent that a pretest market research program cannot fulfill one or more of these objectives, it fails as a significant management decision tool.

Several pre-test-market test methods develop a share-of-market estimate which then must be translated into a volume estimate. In an established market this is not difficult. However, if you are developing a product which is a new category, a share-of-market statistic is not of much help since the objective is to develop a marketing program based on a pre-test-market estimate of volume for the new product. Those pre-test-market techniques which develop volume estimates directly are preferred. Pillsbury has developed a method, "Supertest," to estimate volume levels directly.

Limitations of the Pre-Test-Market Studies

While the wisdom of conducting pre-test-market studies is, I believe, apparent to all, there are limitations. These include the time necessary to do the study, generally the need for finished packaging and advertising, and last, but not least, cost. It is always possible to develop a long list of reasons for not doing something. This is the case in many pre-test-market situations.

It is true that there are situations where pre-test-marketing research is not necessary or desirable. Each situation should be subjected to a cost/benefit analysis. The out-of-pocket costs of the research, the possible lost opportunity represented by the time necessary to test the product, the need to introduce a product into test market for political or other reasons versus the risk of a direct introduction into test market or even nationally must be factored into the decision. No research should be done blindly as part of a necessary step to get a product from concept to national introduction.

Pre-test-market techniques further require certain assumptions regarding awareness and distribution. The Pillsbury technique, "Supertest," has assumptions regarding these two important measures for the success of a new product.[1] To state the obvious, simulated pre-test-market techniques usually have 100 percent awareness among the testers which in most cases will not be attained by any new product regardless of spending, and if an in-store situation is used, the distribution (at least as far as the testing population is concerned) is 100 percent—again, a situation rarely attained by new products. Therefore, modifications of estimates must be made to accommodate the less-than-perfect awareness and distribution which will be obtained nationally. In many cases, these assumptions are based on little fact and much hope. Most product people and most agencies feel that their advertising will generate high levels of awareness, and sales forces rarely say they will not get high levels of distribution. Again, within these constraints there is still a need for pre-test-market volume estimates. Help is available, however. Several advertising agencies have programs to convert the media-plan dollars to gross rating points (GRPs) and then to expected awareness levels. Analysis of awareness levels attained by other new products should be included in the awareness estimates. Expected distribution level can be developed by using levels of current brands, other new products of the company, and competitive levels. These can be adjusted based on the uniqueness of the new product. Ideally a sensitivity analysis using several combinations of projected awareness and distribution should be conducted to alert management to the critical levels necessary for minimal success of the product.

The three pre-test-market methods described in Part III of this book provide an adjustment for distribution and awareness levels based on management judgment concerning the attained level and its impact. Judgmental adjustments are suspect since the desire frequently is to develop high enough volume estimates to allow the product to proceed to test market. Without objective, validated methods of modifying the forecasts for the critical assumptions of distribution and awareness, each of these techniques fails. What is necessary is for manufacturers to develop a history of the awareness and distribution levels attained by their new products at varying

levels of advertising and promotional impact. These can be used to adjust the ongoing awareness and distribution assumptions.

Each of the three models described in Part III meets some of the objectives of a pre-test-market program. (1) All three generate market-share estimates, although none provides an explicit and direct volume estimate for entries into a new-product category. (2) Each provides some elements of a "disaster check." (3) ASSESSOR enables researchers to evaluate variations of a proposed marketing plan by sensitivity analysis. COMP can incorporate variations, for example, in price levels obtained from panel data. Changes in researchers' or managers' judgments about various elements of the proposed marketing program in the LTM model could lead to changes in estimated volume. (4) The attitude measures obtained for ASSESSOR and COMP may provide useful diagnostic insights. A profile of these models and others is included in chapter 7.

Other Methods/Models

While most of the commercial pre-test-market procedures (including the three models presented in Part III) require some kind of simulated purchase decision, there are other models that can be used for pre-test-market assessment. The Hendry model provides a simulation of the marketing program and an estimate of the share the new product will attain within a well-understood and structured market. The use of the Hendry system for pre-test-market analysis assumes that the new product's advertising and packaging effectiveness will equal that of the average product in the market. The Hendry model uses inputs on advertising, pricing, and sales promotion for the new product and all its direct competitors as well as purchase data derived from either a diary panel or a survey. Given such data, it can simulate the likely share and volume the new product will attain. Frequently, the use of this approach as a simulated pre-test-market model is done before the development of a product to decide whether it is worthwhile to enter the market under consideration. Furthermore, as progress is made in the development cycle, new situations emerge which can then be simulated. New insights on the impact of product quality, packaging, advertising effectiveness, and spending levels can be simulated with the model.

An interesting extension of the Hendry model is its use after a decision has been made to go to test market and test markets have been selected. Key to the success of a product in its test market is management expectations. If the projected test-market sales are 3 million units versus expectation of 4 million units, the product is considered a failure. On the other hand, if expectations were 2 million, the product is considered a success. While a brand's national share could be projected to 10 percent, there is probably

no market where it will have that exact share. Some markets will be considerably below 10 percent, and others will be above 10 percent with the national share equaling 10 percent. It is conceivable that a 7 percent share in market *A* would translate to a national share of 10 percent. Conversely, there are markets where a new product can attain considerably above 10 percent, and it would be a mistake to translate that test-market share directly to a national goal. The Hendry model can be of help in this kind of a situation.

If one structures the market in Hendry terms and has an understanding of the competitive shares within the selected test markets, a Hendry simulation can be done to indicate what share of market the new product can attain in that test market. In conjunction with a national Hendry simulation, the expected test-market share becomes the test-market goal. By setting the test-market share objectives to conform with the Hendry simulations, one avoids the problem of trying to explain to management why the 7 percent share attained in test market translates to a 10 percent share nationally. The other side of the coin—why one should not build a national plan on the above-10 percent share—is also avoided. Ideally it would be nice to have test markets for new products with the same marketing plan, but with share expectations above and below the national goal.

Conclusions

With the increasing costs of test market as well as the total competitive frame within a supermarket for new products, it becomes more and more important to know what the new product can be expected to produce before entering a test market. Ultimately it may be possible to go directly from a pre-test-market experiment to a regional or national launch if the pre-test-market prediction ability can be validated. But improvements are necessary on our current techniques. Initially, each company should better understand the pre-test-market techniques they now use and determine how valid the technique is, as they proceed to test market and then to national. Key to improving these techniques are the determination of awareness and distribution, better estimates of repeat, a better fix on appropriate pricing, and the elasticity of advertising and promotion dollars.

As consumer packaged-goods companies, such as Pillsbury, look to their marketing-research departments to reduce the risk of entering test market, we in turn will be looking to outside resources to help us in our pre-test-marketing programs. However, cooperation between consumer packaged-goods companies should be encouraged. The professionalism of the function can be enhanced as each of us proceeds with our own models to determine the volume expected for a new product. As we gain experience

with syndicated services, sharing of information will enhance the predictability of pre-test-marketing measures, while at the same time providing more assurance to management that all that can be done is being done by the marketing-research professional.

Note

1. "Supertest" takes place in supermarkets where qualified respondents are shown TV commercials of the test product and other products and then given discount coupons good for the day to stimulate buying of these products. After the in-store portion of the Supertest, there are follow-up telephone interviews of those who redeemed coupons to obtain after-use reaction to the test product and some repeat information. In addition, test products are left in the test stores for four weeks after the in-store portion of the test to monitor the store sales and repeat purchases in a four-week call-back. The consumer-dynamics parameters are estimated by using a set of empirically derived models based on an in-house developed and maintained data base. Similarly, the volume estimation uses a Pillsbury-developed depth-of-repeat model. The volume projections are conditional to the achievements of level of awareness and distribution.

Part IV
Test-Market-Based New-Product Forecasting Methods

The performance evaluation of the new product *often* centers on two major aspects: the prediction of the sales, profit, and market share of a new product over time and the evaluation of the feasibility and performance of alternative marketing strategies.

The test-market stage is usually the first time in the new-product development process that *all* the relevant marketing variables can be tested *together*. Alternative marketing strategies based on different combinations of packaging, pricing, advertising, trade and consumer promotional and distribution mixes can be tested under realistic market conditions for a reasonable length of time.

To assess the performance of the product and its associated marketing mix, most test markets rely heavily on *store audits* which measure primarily the product's sales volume and market share. These measures often are supplemented by a number of waves of consumer surveys, which measure the levels of and changes in consumers' awareness of and attitudes toward the test product, the incidence of trial, and, most importantly, the repeat-purchase pattern and volume. Assessing the repeat-purchase probabilities of the new product and level of purchase is a critical component of the new product's performance evaluation and requires the establishment of a consumer panel in the test-market areas.

Whatever the sources of the data used, the primary output of test market should be projected sales, profit, and share conditional on the alternative marketing strategies under evaluation. To obtain such output, a prediction model has to be established to translate the actual results of the test market into projected annual national sales, profits, and market share. The design of a test market, therefore, should take into account the data that would be required as input to the new-product forecasting model chosen by the firm.

Chapter 12 by Narasimhan and Sen provides a critical evaluation of nine models which have been developed to evaluate test-market results. More specifically, the chapter compares these models on seven dimensions: model objective, stages modeled (awareness, trial, repeat, and so on), consideration of marketing-mix variables, model complexity, type of sales data required, diagnostic aid and planning, and degree of commercial acceptance. Two of the nine models reviewed in chapter 12 and developed by Parfitt and Collins, and Blattberg and Golanty, respectively, are included as

illustrative models (chapters 13 and 16, respectively). In addition, the underlying structure of repeat-buying models is examined in the chapter by Kalwani and Silk (chapter 15).

One of the earlier test-market-data-based models, also included in this part, chapter 14, was developed by Claycamp and Liddy for N.W. Ayer. Chapter 17 by Dodson, in addition to discussing the applications of test-market models in general, provides further details on this model and discusses the utilization of this model for pre-test-market forecasting.

Additional Readings

Wind, Yoram. *Product Policy: Concepts, Methods, and Strategy.* Reading, Mass.: Addison-Wesley Publishing Company, 1981, chap. 14.

"Market Testing Consumer Products: A Symposium." *Experiences in Marketing Management,* No. 12. New York: The Conference Board, 1967.

12 Test-Market Models for New-Product Introduction

Chakravarthi Narasimhan and
Subrata K. Sen

New-Product Introduction Process

The introduction of a new product is preceded by a series of important steps, as indicated in figure 12–1. At the end of each step, management evaluates the likelihood of success of the new product and makes a go or no go decision. The first step in the process consists of the definition of the market, that is, the definition of the product category in which the new product is to be introduced (see Day, Shocker, and Srivastava 1979). The second step, idea generation and screening, refers to the generation of alternative new-product ideas and some preliminary evaluation of the viability of the proposed ideas. The ideas that survive the screening step are formulated as new-product concepts. A product concept is typically defined in terms of a specific bundle of product attributes. The concept-evaluation stage consists of a detailed, formal analysis of each product concept using techniques such as multivariate statistical procedures, conjoint analysis, and mathematical programming. Stages two and three of figure 12–1 are reviewed in detail by Shocker and Srinivasan (1979) while the methodological and measurement issues in the concept-evaluation stage are discussed by Hauser and Urban (1977).

The output of the concept-evaluation stage consists of a very small number of concepts (perhaps one or two) which have to be analyzed in more detail. The next step is to manufacture an actual prototype of the product concept which, so far, has been represented only as a bundle of product attributes. Until recently, the product would have been introduced into a set of test markets in order to obtain a sales prediction. However, test marketing is an expensive way to find out that a product is likely to be unsuccessful. Additionally, it reveals the company's new-product strategy to its competitors. Hence, there has been an increasing trend toward pre-test-market evaluation (see figure 12–1), which is essentially a laboratory simulation of the test-market process (see Tauber 1977 for a review of pre-test-market models). Silk and Urban (1978) provide a comprehensive model for the analysis of pre-test-market data[1] while the approach presented by Claycamp and Liddy (1969) is another way of predicting a new product's sales without actually introducing it in a test market.

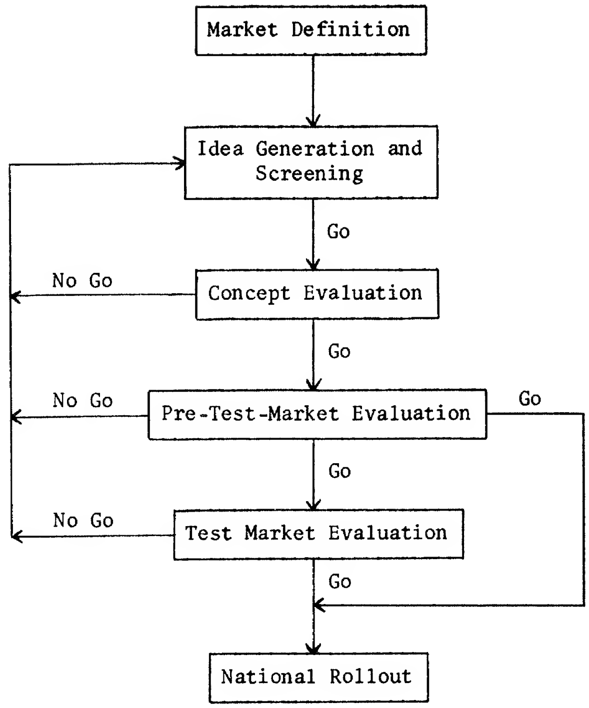


Figure 12-1. The New-Product Introduction Process

If the product appears to be successful based on a pre-test-market evaluation, it is either introduced nationally or, more frequently, introduced into a set of test markets (see figure 12-1). The test markets are used for two purposes: validation of the sales forecasts made at the pre-test-market stage and evaluation of alternative marketing mixes for the new product (see Achenbaum 1974 and Klompmaker, Hughes, and Haley 1976).

The purpose of this chapter is to provide a critical review of several models which have been proposed to evaluate test-market results. Some of the models reviewed focus only on sales prediction while others (generally, the newer models) emphasize both sales prediction and the impact of marketing-mix variables. The nine models reviewed here are listed in table 12-1. All the models deal with a new *brand* in an inexpensive, frequently purchased product category. These particular models were chosen for review because they are the most frequently cited test-market models in the marketing literature.[2]

Table 12–1
Classification of Test-Market Models

		Modeling Aspects					
Model	*Model Objective*	*Level of Model Complexity*	*Quality of Modeling*	*Type of Sales Data Required*	*Diagnostics*	*Number of Applications Reported*[a]	*Degree of Commercial Acceptance*
Fourt and Woodlock (1960)	Sales	Low	Unsatis-factory	Panel	Low	7	Low
Parfitt and Collins (1968)	Brand share	Low	Unsatis-factory	Panel	Low	25	High
STEAM (Massy 1969)	Sales	High	Unsatis-factory	Panel	Low	2	Low
SPRINTER (Urban 1970)	Sales	High	Good	Panel store audit	High	1	Low
Eskin (1973)	Sales	Medium	Unsatis-factory	Panel	Low	6	Medium
Nakanishi (1973)	Sales	High	Unsatis-factory	Panel	Medium	1	Low
NEWS (BBDO International, Inc. (undated)	Sales/Profit-ability analysis	Medium	Don't know	Panel	High	33	High
NEWPROD (Assmus 1975)	Brand share	Medium	Satisfactory	Survey	Medium	7	Low
TRACKER (Blattberg and Golanty 1978)	Sales	Medium	Good	Survey	High	11	High

[a]At the time of publication.

Dimensions for Model Evaluation

In this section we briefly describe the dimensions on which the various models are evaluated. The next section provides an evaluation of each model in terms of these dimensions.

Model Objective

The first dimension deals with the criterion variable that must be estimated from test-market data. Year-end sales forecast, market-share forecast, and some measure of profitability are likely candidates for the criterion measure. Profitability is clearly the best criterion, but unlike the other criteria, it requires cost as well as sales data.

Modeling Aspects

There are three main aspects to this dimension: consideration of the different stages of new-product acceptance, the mathematical structure of the models for each stage, and issues related to the estimation of the model's parameters.

Stages Considered. The three main stages through which a consumer moves in the process of becoming a customer of a new brand are indicated by the square blocks in figure 12–2. The three stages consist of the awareness class, the trier class, and the repeater class. The awareness class refers to consumers who have become aware of the new brand through the firm's promotional activities, the trier class refers to consumers who have made one purchase of the new brand, and the repeater class refers to consumers who have purchased the new brand more than once.[3]

Regarding the modeling of these stages, the first task is to model the link between the awareness and trial classes explicitly. In particular, two aspects of this link should be considered:

1. Consumer segments which become aware of the new brand from different sources (for example, advertising, coupons, or free samples) may differ in their trial probability.
2. The trial probability may also differ between those who become aware in the current period and those who became aware in some earlier period.

Similarly, one should be concerned with different trier classes in terms of their *time of entry* into the trier class for the new brand. Empirical evi-

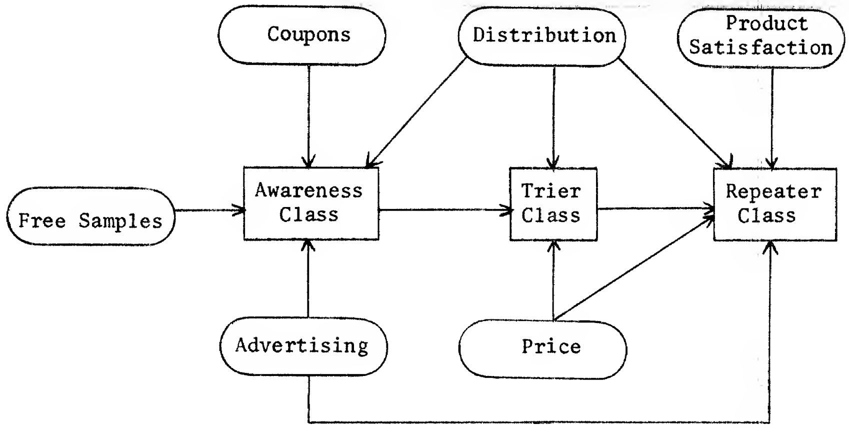


Figure 12-2. Framework for Test-Market Models

dence exists (see Parfitt and Collins 1968) that early triers and late triers exhibit different repeat-purchase probabilities.

One should also distinguish between different levels of repeaters. In other words, consumers who have bought once and are repeating for the first time may be different in their purchase probability from those who have already repeat-purchased the new brand, say, five times.[4]

If the issues discussed above are not considered explicitly, the model's predictions are likely to be inaccurate.

Mathematical Models for Each Stage. The main issue here is the extent to which the models consider the effects of marketing-mix variables. This issue is examined in figure 12-2 where the oval blocks represent the marketing variables that affect a consumer's movement from one stage to another. Figure 12-2 indicates that free samples, coupons, advertising, and distribution make the consumer aware of the new brand. The consumer's awareness level, along with the new brand's price and distribution, determine whether the consumer will try the product. The trial experience, positive or negative, along with distribution, price, and advertising, will affect the consumer's decision to repeat-purchase the brand.[5]

It is also important to consider whether the *appropriate* marketing-mix variables are incorporated in the models for each stage (the appropriate marketing-mix variables are indicated in figure 12-2). For example, a model of trial probability that includes distribution but not price will be viewed as a model that does not include the appropriate marketing-mix variables.[6]

Estimation Issues. Regarding estimation, we examine whether the model description reports specific methods of estimating the model's parameters and values of the estimated parameters.

Sales-Data Requirements

To gather sales data in the test market, one could use either diary panel data or survey data. Panel data are likely to be more accurate, but are more expensive than survey data. The survey method also permits the researcher to design questions to infer consumers' awareness, repeat-purchase probability, and so forth, an advantage not obtainable from diary panel data. This is important in generating diagnostic information from the model.

Diagnostic Aid

Apart from meeting the objective of providing a sales forecast or a forecast of brand share, a test-market model must aid the user in generating useful diagnostic information on the effectiveness of the various marketing-mix variables. If the sales of a well-designed new product are low, management would like to know why the product is not performing better. For example, the awareness level for the new brand could be low, suggesting that the firm's advertising, couponing, or distribution of free samples has been ineffective. Or the trial rate may be low, suggesting probable bottlenecks in distribution or too high a price. Whatever the problem, it is necessary to isolate and identify it so that suitable corrective action can be taken. Even for a successful introduction, the firm is interested in knowing the relative effectiveness of its marketing instruments.[7]

Degree of Commercial Acceptance

The final dimension for evaluating the various test-market models is the degree to which these models have been used commerically. Clearly, this is a difficult factor to assess. However, because it is an important issue for managers, we have tried to assess it in two ways. First, we indicate the number of applications reported in the model's description. Papers dealing with models which have been frequently applied are likely to report a larger number of applications. Second, since the number of reported applications may not reflect the actual degree of use of the model, we provide estimates of the degree of commercial use of each model. These estimates are based primarily on subjective information obtained from a variety of sources such as the

developers of the models and executives of diary panel firms, for example, the Market Research Corporation of America.[8]

Critical Evaluation of the Models

In this section we provide a critical evaluation of the nine models in terms of the dimensions described in the preceding section. Our evaluation consists of a comparison of the different models on each evaluative dimension. The highlights of this comparison are presented in table 12–1.

Model Objective

Table 12–1 reports the different criterion measures (for example, sales, brand share) that are forecasted by the various models.[9] Only the NEWS model makes an explicit estimate of the new brand's profitability. The other models predict the sales or the market share of the new brand.

Modeling Aspects

This dimension deals with a series of criteria that relate to the structure of the model and the mathematical and estimation details provided in each model's write-up. Our evaluation of the nine models on these criteria is summarized in table 12–2. A "yes" indicates that the model considers the particular issue or reports the particular model or parameter estimate. A "no" indicates that it does not. "Partial" indicates that the issue is only partially dealt with. For example, the Parfitt-Collins model describes a specific mathematical model for trial but does not provide a specific model for repeat purchases. Consequently, table 12–2 shows an entry of "partial" for the Parfitt-Collins model in the column titled "Report of Specific Models."

Table 12–2 provides only a summary evaluation of the models on the modeling-aspects dimension. Readers interested in a more detailed description of each model are referred to Appendix 12A. The remainder of this section provides a critical evaluation of each model in terms of the criteria listed in table 12–2.

Fourt and Woodlock (1960). This is a very simple model, but it suffers from the following shortcomings:

1. The awareness stage is not modeled. Thus, the effect of different ways of becoming aware and of the time of becoming aware on the probability of trial cannot be studied.

Table 12-2
Modeling Details

		Details Regarding Each Stage				*Mathematical Models for Each Stage*			*Estimation Issues*	
		Awareness		*Trial*	*Repeat*					
Model	*Stages Modeled*	*Different Ways of Becoming Aware and Trial Probability*	*Time of Becoming Aware and Trial Probability*	*Time of Becoming a Trier and Repeat-Purchase Probability*	*Depth of Repeat and Repeat-Purchase Probability*	*Report of Specific Models*	*Marketing-Mix Variables Considered*	*Inclusion of Appropriate Marketing-Mix Variables*	*Report of Estimation Methods*	*Report of Parameter Estimates*
Fourt and Woodlock (1960)	Trial, repeat	—	—	No	Yes	Yes	None	—	No	No
Parfitt and Collins (1968)	Trial, repeat	—	—	Yes	No	Partial	None	—	Partial	No
STEAM (Massy 1969)	Trial, repeat	—	—	Yes	Yes	Yes	None	—	Yes	No
SPRINTER (Urban 1970)	Awareness, trial repeat	Yes	No	No	Yes	Yes	Advertising, price, promotion, distribution	Yes	Yes	No
Eskin (1973)	Trial, repeat	—	—	Yes	Yes	Yes	None	—	Yes	Yes
Nakanishi (1973)	Trial, repeat	—	—	Yes	No	Yes	Advertising, promotion	No	Yes	Yes
NEWS (BBDO International Inc. undated)	Awareness, trial, repeat	Don't know	Don't know	Don't know	Don't know	No	Advertising, price, promotion, distribution	Yes	No	No
NEWPROD (Assmus 1975)	Awareness, trial, repeat	Yes	Yes	No	No	Partial	Advertising, promotion	No	No	No
TRACKER (Blattberg and Golanty 1978)	Awareness trial, repeat	No	Yes	No	Yes	Yes	Advertising, price	Yes	Yes	Yes

2. The time of entry into the trial class and its impact on repeat-purchase probability are not considered.
3. Marketing-mix variables are totally ignored, making it extremely difficult for management to assess the effectiveness of the marketing tools under its control.
4. Accurate calculation of the various repeat-purchase probabilities requires very large samples.
5. Since estimation details and parameter estimates are not reported, it is difficult to assess the quality of the model's predictions.

Parfitt and Collins (1968). The main contributions of the model are:

1. It is a simple model that appears to predict a new brand's market share accurately. In addition, it provides some useful diagnostic information if the brand's predicted market share S is very low. If S is low because the penetration P is low while the repeat-purchase rate R is at an acceptable level, it means that the marketing mix for the product should be altered to induce greater trial. Once consumers try the product, they seem to like it enough to repurchase it an an acceptable rate. If, on the other hand, P is high but R is low, the product should be withdrawn or reformulated because triers of the brand do not appear to like it, as evidenced by the low repurchase rates.
2. It was clearly demonstrated that the repeat-purchase rates were different for customers entering at different times after the new product was introduced. This is important in estimating the equilibrium repeat-purchase rate R.
3. The authors show that promotions such as temporary price cuts guarantee increased penetration but not necessarily a higher repeat-purchase rate. The increase in repeat-purchase rate may be nominal, which may not warrant the cost of the promotion.

The model's shortcomings consist of:

1. The awareness stage is not modeled.
2. The impact of depth of repeat on repeat-purchase probability is not considered.
3. The absence of marketing-mix variables makes it impossible to assess the effects of advertising, sampling, or couponing.
4. The subjective estimation of the equilibrium repeat-purchase rate may create problems. Not only is the method ad hoc, but also it does not provide us with an estimate of the forecast error.
5. The model has been criticized in terms of the small sample sizes of the panels used to estimate the model's parameters. Shoemaker and Staelin

(1976) show that the panel sizes that are typically used to estimate the parameters of the model are often too small to allow us to place much confidence on the forecasted brand share. This occurs mainly because the equilibrium repeat-purchase rate cannot be estimated precisely.

STEAM (Massy 1969). This model is extremely complex and suffers from the following problems:

1. It does not model the awareness stage.
2. It does not consider the effects of marketing-mix variables.[10]
3. Large sample sizes are required to estimate the model's parameters. The application reported uses a panel of about 6,000 households. Given that sufficient observations are required for each depth of trial class in order to obtain reliable estimates, one would suspect that it is necessary to have extremely large samples sizes. This clearly increases the cost of implementing the model.

SPRINTER (Urban 1970). On the positive side, SPRINTER traces out the consumer adoption process carefully and considers the effects of marketing-mix variables explicitly. However, the model can be criticized on the following grounds:

1. It does not consider the time of becoming aware and its impact on trial probability.
2. Time of entry into the trial class and its impact on repeat-purchase probability are not modeled.
3. Because of the complexity of the model (it consists of approximately 500 equations), data requirements are very formidable.

Eskin (1973). This model is positioned between the Parfitt-Collins model and the STEAM model. Thus, it is more complex than the former (that is, the purchase process is modeled in greater detail), but is computationally less burdensome than the latter. The model suffers from the following shortcomings: It does not model the awareness stage, and the effect of marketing-mix variables is not modeled explicitly. Thus, any kind of diagnostic information on the effectiveness of these tools is not obtainable from the model.

Nakanishi (1973). The problems with this model are as follows:

1. Even though Nakanishi mentions a "pretrial" stage when the consumer becomes aware of the new product, the awareness stage is not modeled explicitly.

2. The impact of depth of repeat on repeat-purchase probability is not considered.
3. The trial and repeat models do incorporate the effects of marketing-mix variables. However, the trial model does not consider the *appropriate* marketing-mix variables since it ignores the effect of relative price on trial.
4. The fit obtained for the repeat-purchase model is extremely poor (R^2 0.1), leaving open the question of the appropriateness of the model and/or estimation procedure.

NEWS (BBDO International, Inc., undated). It is difficult to provide a critical evaluation of NEWS because we do not have a detailed description of the model (see the discussion of NEWS in Appendix 12A). NEWS does appear to deal with the relevant stages of new-product acceptance and explicitly considers the appropriate marketing-mix variables. However, we do not have the information to comment on the quality of the modeling.

NEWPROD (Assmus 1975). On the positive side, the model does examine the effect of marketing-mix variables in determining consumer acceptance of a new brand. However, NEWPROD can be criticized on the following grounds:

1. The impact of time of entry into the trier class on repeat-purchase probability is not modeled.
2. The effect of depth of repeat on repeat-purchase probability is ignored.
3. Relative price is not included in the trial model.
4. In general, the details of the model are provided in a very sketchy manner, making it difficult to comment on the quality of the modeling.

TRACKER (Blattberg and Golanty 1978). This model has the following advantages:

1. Since marketing-mix variables are incorporated in the model, TRACKER generates useful diagnostic information about the new product.
2. Since it uses only survey data, it is considerably cheaper than models that use diary panel data. Yet, TRACKER provides forecasts of comparable accuracy.

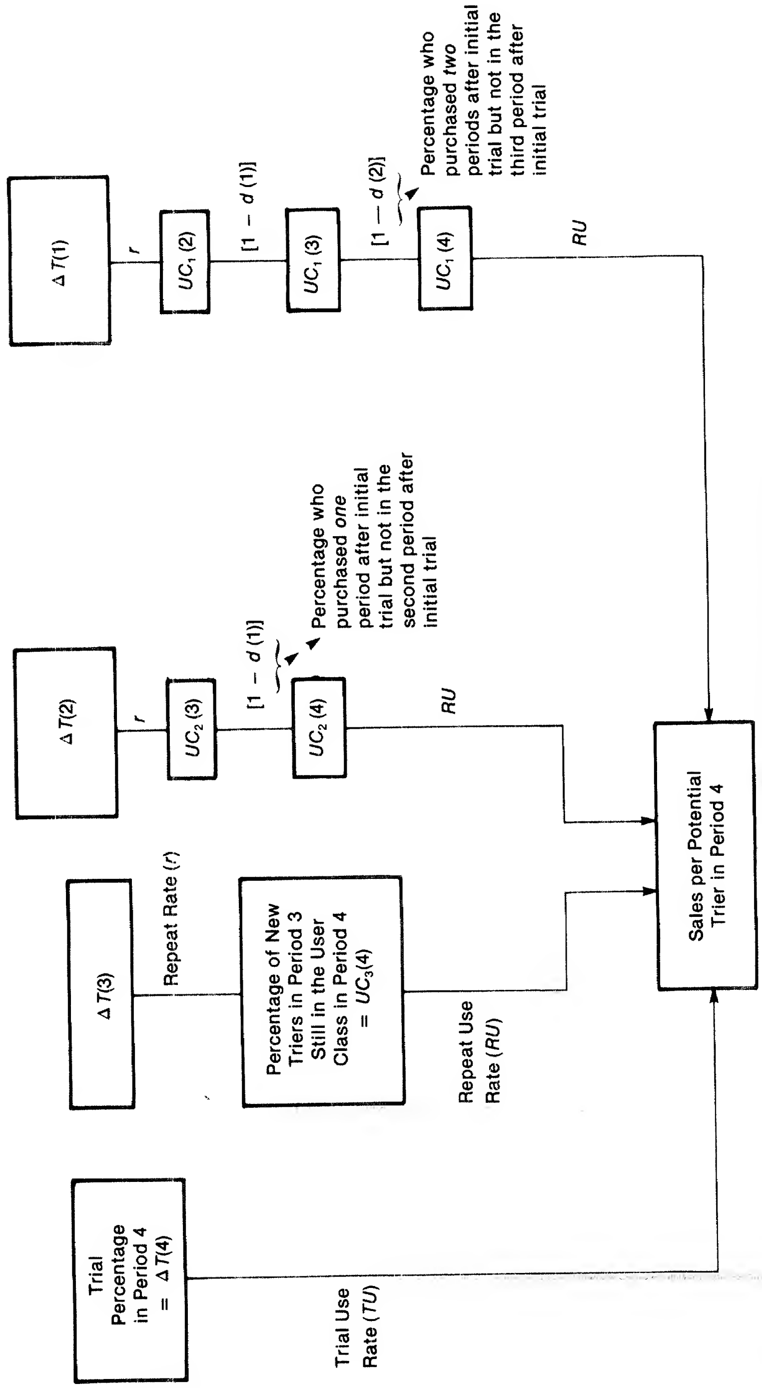
Some of the weaknesses of the model consist of the following:

1. Although TRACKER models all three stages of the adoption process, table 12-2 indicates that it does not consider all the details of the awareness and trial stages.

2. The model does not explicitly consider the distribution of free samples and coupons. The authors state that when a new product is sampled heavily, the effect of sampling can be taken into account by making suitable modifications to the model. However, these modifications are not described.
3. The assumption that all brands in a product class are equally responsive to marketing instruments is unappealing. For example, in determining the parameters of the awareness model, all brands are pooled and a single set of response parameters is estimated. Although data limitations force the authors to proceed in this manner, some of the unsatisfactory results in the awareness model's predictions may stem from the untenable assumption of parameter homogeneity across brands. If one looks at the "Absolute Error" column in table 16-4, one observes some extremely large errors, even though we are looking at the same data that were used to estimate the model's parameters.
4. In the projection model, most of the estimates are subjectively determined. Further, the assumption that the repeat rate r does not vary with time of entry into the market is contrary to empirical evidence (Parfitt and Collins 1968). Since the model seems to predict well, it appears that it is not very sensitive to such parameters except the $d(i)$'s (which are defined in figure 12-3).

The model evaluations in table 12-2 are summarized in table 12-1 in terms of the two columns labeled "Level of Model Complexity" and "Quality of Modeling." Level of model complexity refers to the complexity of the model structure (the number of variables, the number of equations, and so on) as well as the complexity of the analytical and estimation techniques used. Our overall evaluation of each model's complexity is indicated in table 12-1 in terms of three levels: low, medium, and high. The Fourt-Woodlock and Parfitt-Collins models are simple models while the STEAM, SPRINTER, and Nakanishi models are very complex. The other four models have a medium level of complexity.

The quality-of-modeling criterion is based on two main aspects: the degree to which the various stages of the new-product adoption process are modeled and the extent to which appropriate marketing-mix variables are incorporated in these models. If *all three stages* are modeled and *appropriate* marketing-mix variables are considered for each model, the quality-of-modeling rating is good in table 12-1. The rating is satisfactory if the model considers all three stages and incorporates the effects of marketing-mix variables, even though some of the marketing variables are not considered in an appropriate manner. The rating in table 12-1 on this criterion is unsatisfactory if the model does not consider all three stages *or* totally ignores the effect of marketing-mix variables. Based on these criteria, the quality of the



$$TS(4) = \underbrace{TU\ \Delta T(4)}_{\text{new triers}} + \underbrace{RU\,[UC_3(4)]}_{\text{first repeaters}} + \underbrace{RU\,[UC_2(4)]}_{\text{second repeaters}} + \underbrace{RU\,[UC_1(4)]}_{\text{third repeaters}}$$

Figure 12–3. TRACKER: Sales-Projection Model for Period 4

modeling is good only for SPRINTER and TRACKER. It is satisfactory for NEWPROD and is unsatisfactory for the remaining models except NEWS, which cannot be evaluated on this criterion because of the lack of information.

Type of Sales Data Required

As indicated in table 12–1, most of the models require diary panel sales data. SPRINTER requires store-audit data in addition to panel data. Only the NEWPROD and TRACKER models use survey data. Consequently, they are likely to be cheaper than the other models listed in table 12–1.

Diagnostic Aid

To be an effective diagnostic aid, a model must explicitly incorporate the effects of marketing-mix variables. Our evaluation of the diagnostic capabilities of each model is summarized in table 12–1 in terms of three levels: low, medium, and high. NEWS and TRACKER clearly have high diagnostic capabilities as does SPRINTER. NEWPROD and Nakanishi's model have some diagnostic capabilities while the four remaining models are weak on this dimension.

Degree of Commercial Acceptance

On this dimension, table 12–1 first indicates the number of applications reported for each model. We attempted to determine how many of these applications consisted of predictions which were later validated by actual sales. However, with some exceptions (discussed below), this information was difficult to determine from the model descriptions.

It was also difficult to determine the product categories for which the various models have been used. Because of confidentiality requirements, many authors (especially the authors of commercial models such as NEWS and TRACKER) tend not to report the product categories for which their model was applied. Thus, with very few exceptions, all we know is that the model applications were for inexpensive, frequently purchased products. Whatever product-category information is available is mentioned below.

The Parfitt-Collins model has essentially superceded the Fourt-Woodlock model. According to executives of the Market Research Corporation of America (MRCA), the Parfitt-Collins model has gained wide acceptance

in the United States in addition to being used extensively in Britain. The Parfitt-Collins model has been applied to product categories such as toothpaste, dishwashing liquid, toilet tissue, butter, instant coffee, and floor polish. Most of the applications reported in Parfitt and Collins (1968) appear to consist of predictions that were later validated by actual sales. The Fourt-Woodlock model has been applied to products such as cake mix, cereals, margarine, detergents, pet foods, canned fruits, and scouring pads.

The development of STEAM was funded by diary-panel companies such as MRCA. However, there is no evidence that it is currently being used by any company. Nor has the full-blown SPRINTER model been used very much. However, there have been some applications of a simplified version of the model. Eskin's model was developed for the Pillsbury Company and is presumably being used by Pillsbury. The extent of its use, however, is not clear. The Nakanishi model was applied to the tomato-catsup product category by the author. However, there is no evidence that this model is used commercially. The NEWPROD model was developed for a specific unnamed corporation, but it is not clear to what extent the model is used by the company.

The NEWS brochure (BBDO International, Inc., undated) states that the model has been requested by thirty-seven advertisers for nearly 100 products, and validation data for thirty-three of these products are provided in the brochure. NEWS appears to have gained wide acceptance with many of BBDO's clients. Finally, the TRACKER model was developed for the Leo Burnett advertising agency. As reported in Blattberg and Golanty (1978), there is ample evidence of its use by clients of Leo Burnett.

Our overall evaluation of the degree of commercial acceptance of these models is summarized in the last column of table 12-1 in terms of three levels of acceptance: low, medium, and high. NEWS, TRACKER, and the Parfitt-Collins model appear to have gained the greatest commercial acceptance.

Overall Evaluation

Overall, the TRACKER and NEWS models appear to be the best of the test-market models reviewed. Both models are complete in terms of the stages modeled and the inclusion of marketing-mix variables. Yet, they are not so complex that they are difficult and costly to implement. Both models provide good diagnostics, predict well, and appear to enjoy a high degree of commercial acceptance. TRACKER has an edge over NEWS for two main reasons. First, TRACKER uses survey data and is therefore likely to be cheaper to implement than NEWS, which requires diary-panel sales data. Next, the quality of the modeling in TRACKER is a known quantity and

has been judged as being good. In comparison, we have very little information on the quality of the modeling in NEWS. Consequently, we are uncertain about an important aspect of NEWS, making it difficult to give it an unqualified recommendation.

Besides TRACKER and NEWS, the Parfitt-Collins model also merits consideration. It is very simple to understand and explain to managers and appears to provide accurate predictions. However, the Parfitt-Collins model has the important drawback of not modeling the effects of marketing-mix variables explicitly.

Conclusions and Future Directions of Research

In this chapter we have reviewed a number of test-market models which have appeared in the literature. These models vary markedly in terms of model design, complexity, and data requirements. Many of the models claim excellent fit to the data. This makes one wonder about the value of building very sophisticated models and using powerful estimation techniques. It would be interesting to apply a subset of these models to the same data set(s) and examine their comparative performance.

As experience with these models increases, both in the number of applications and in the number of different product classes where the models have been applied, we believe that managers will gain sufficient insights into the ranges of the models' parameter values. This should enable the manager to use these models for planning purposes in a pre-test-market sense for future new products. In other words, the model could be used to simulate the sales that would be generated by alternative marketing programs for the new brand. Thus, the simulations would enable the firm to choose the best marketing mix for the brand. This "planning" aspect is stressed in the publications describing both NEWS and TRACKER. In addition, a pre-test-market version of the test-market model by Eskin (1973) is currently being developed (see Eskin and Malec 1976 for an early write-up of this model).

The state of the art for building test-market models appears to be in an advanced state for frequently purchased products. On the other hand, the evaluation of test-market data for infrequently purchased durables has not been tackled by marketing-model builders. For this purpose, the first-purchase diffusion models (Mahajan and Muller 1979) or the ideas behind the diffusion models which incorporate repeat purchasing (see, for example, Dodson and Muller 1978) are likely to be useful. Of course, such models should incorporate the effects of marketing-mix variables such as price and advertising.

Notes

1. Elrick and Lavidge, Yankelovich, and other marketing-research companies also have pre-test-market models of their own. Elrick and Lavidge's model, COMP, is described in some detail in this book (chapter 10), which also contains an outline of the Yankelovich model (chapter 9).

2. New-product diffusion models are not reviewed because most diffusion models deal only with the *first* purchase of the new product (see Mahajan and Muller 1979). Consequently, they are primarily applicable for durables and are not applicable for frequently purchased products where repeat purchasing is very important. At best, they might be useful in modeling the first stage (the penetration stage) of the test-market models reviewed here (see Midgley 1976, for example). Even for this purpose, their applicability is not entirely clear since the first-purchase diffusion models typically refer to a *product class* and not to a *brand* in a product category. Some diffusion models (most notably Dodson and Muller 1978) have tried to model the repeat-purchase phenomenon that characterizes the purchase of inexpensive, frequently purchased products. However, these diffusion models focus primarily on issues such as word-of-mouth recommendations which, though very important for the purchase of durables, are unlikely to be important for inexpensive, frequently purchased products. The consumer can purchase and experience several alternative brands in such product classes at a relatively low cost. This would certainly appear to be a more efficient way of acquiring information about such products than seeking word-of-mouth recommendations. It appears, therefore, that diffusion models are really not appropriate for the type of products that are dealt with by most test-market models.

3. It might be argued that the models should also explicitly consider changes in the sales of the *product class*. However, if the new brand is introduced in an established product class, one could probably ignore the effect on primary demand and assume that the potential trier class for the new brand is fixed.

4. Although it is theoretically desirable to model different classes of repeaters separately, some recent empirical work by Kalwani and Silk (1980) appears to indicate that it is adequate to explicitly model the first repeaters only.

5. In actuality, all the marketing-mix variables listed in figure 12–2 will affect each of the three stages in varying degrees. We have attempted to isolate the most important effects in figure 12–2. In addition, other factors could influence the consumer to move from one stage to another. For example, a consumer who is aware of the new product may choose to buy it out of curiosity or variety seeking without being unduly affected by the prod-

uct's price. We have elected to omit such factors because they are difficult to model and none of the models reviewed explicitly take such factors into account.

6. In the above discussion we have highlighted some key issues that should be considered while building a test-market model. However, we are aware that increased sophistication at the modeling stage is not costless if the model is to be adopted and implemented by marketing managers. A model which is theoretically sound may be intractable in terms of its mathematical formulation, data requirements, and estimation procedures. We are simply stating that the model builder should be aware of these issues and that any tradeoffs that become necessary should be made with sound managerial judgment regarding costs, ease of estimation, and model tractability.

7. To be an effective diagnostic aid, a model must explicitly incorporate the effect of marketing-mix variables in modeling the various stages of new-product acceptance.

8. A more systematic approach to determining the practical usefulness of a model is published in Larreche and Montgomery (1977) who evaluated three of the models we examine: SPRINTER, STEAM, and NEWS, on the criterion of managerial acceptability.

9. Table 12–1 lists only the ultimate measure of interest, ignoring intermediate measures forecasted by some of the models. For example, models such as NEWS, SPRINTER, and TRACKER can provide forecasts of intermediate measures such as brand awareness and trial probability.

10. However, Massy, Montgomery, and Morrison (1970, pp. 428–435) briefly discuss how marketing-mix variables can be incorporated into STEAM.

References

Achenbaum, Alvin. 1974. "Market Testing: Using the Marketplace as a Laboratory." In *Handbook of Marketing Research,* edited by Robert Ferber. New York: McGraw-Hill, pp. 4:31–4:54.

Ahl, David H. 1970. "New Product Forecasting Using Consumer Panels." *Journal of Marketing Research* 7 (May):160–167.

Assmus, Gert. 1975. "NEWPROD: The Design and Implementation of a New Product Model." *Journal of Marketing* 39 (January):16–23.

BBDO International Inc., Undated. "How to Predict New-Product Success." Unpublished brochure.

Blattberg, Robert, and Golanty, John. 1978. "TRACKER: An Early Test-Market Forecasting and Diagnostic Model for New Product Planning." *Journal of Marketing Research* 15 (May):192–202.

Claycamp, Henry J., and Liddy, Lucien E. 1969. "Prediction of New Product Performance: An Analytical Approach." *Journal of Marketing Research* 6 (November):414–420.

Day, George S.; Shocker, Allan D.; and Srivastava, Rajendra K. 1979. "Customer-Oriented Approaches to Identifying Product Markets." *Journal of Marketing* 43 (Fall):8–19.

Dodson, Jr., Joe A., and Muller, Eitan 1978. "Models of New Product Diffusion through Advertising and Word-of-Mouth." *Management Science* 24 (November):1568–1578.

Eskin, Gerald J. 1973. "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model." *Journal of Marketing Research* 10 (May):115–129.

Eskin, Gerald J., and Malec, John. 1976. "A Model for Estimating Sales Potential prior to the Test Market." *Proceedings of the American Marketing Association.* Chicago: American Marketing Association, pp. 220–233.

Fourt, Louis A., and Woodlock, Joseph W. 1960. "Early Prediction of Market Success for New Grocery Products." *Journal of Marketing* 25 (October):31–38.

Hauser, John R., and Urban, Glen L. 1977. "A Normative Methodology for Modeling Consumer Response to Innovation." *Operations Research* 25 (July–August):579–619.

Kalwani, Manohar U., and Silk, Alvin J. 1980. "Structure of Repeat Buying for New Packaged Goods." *Journal of Marketing Research* 17 (August):315–322.

Klompmaker, Jay E.; Hughes, G. David; and Haley, Russell I. 1976. "Test Marketing in New Product Development." *Harvard Business Review* 54 (May/June):128–138.

Larreche, Jean-Claude, and Montgomery, David B. 1977. "A Framework for the Comparison of Marketing Models: A Delphi Study." *Journal of Marketing Research* 14 (November):487–498.

Learner, David B. 1968. "Profit Maximization through New Product Marketing Planning and Control." In *Applications of the Sciences in Marketing Management,* edited by Frank M. Bass, Charles W. King, and Edgar A. Pessemier. New York: Wiley, pp. 151–167.

Light, Lawrence, and Pringle, Lewis. 1970. "New Product Forecasting Using Recursive Regression." In *Research in Consumer Behavior,* edited by David T. Kollat, Roger D. Blackwell, and James F. Engel. New York: Holt, Rinehart and Winston, pp. 702–709.

Mahajan, Vijay, and Muller, Eitan. 1979. "Innovation Diffusion and New Product Growth Models in Marketing." *Journal of Marketing* 43 (Fall):55–68.

Massy, William F. 1969. "Forecasting the Demand for New Convenience Products." *Journal of Marketing Research* 6 (November):405–412.

Massy, William F.; Montgomery, David B.; and Morrison, Donald G. 1970. *Stochastic Models of Buying Behavior.* Cambridge, Mass.: M.I.T. Press.

Midgley, David F. 1976. "A Simple Mathematical Theory of Innovative Behavior." *Journal of Consumer Research* 3 (1976):31–41.

Nakanishi, Masao. 1973. "Advertising and Promotion Effects on Consumer Response to New Products." *Journal of Marketing Research* 10 (August):242–249.

Parfitt, J.H., and Collins, B.J.K. 1968. "Use of Consumer Panels for Brand Share Prediction." *Journal of Marketing Research* 5 (May):131–145.

Shocker, Allan D., and Srinivasan, V. 1979. "Multiattribute Approaches for Product Concept Evaluation and Generation: A Critical Review." *Journal of Marketing Research* 16 (May):159–180.

Shoemaker, Robert, and Staelin, Richard. 1976. "The Effects of Sampling Variation on Sales Forecasts for New Consumer Products." *Journal of Marketing Research* 13 (May):138–143.

Silk, Alvin J., and Urban, Glen L. 1978. "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology." *Journal of Marketing Research* 15 (May):171–191.

Tauber, Edward M. 1977. "Forecasting Sale prior to Test Market." *Journal of Marketing* 41 (January):80–84.

Urban, Glen L. 1970. "Sprinter Mod III: A Model for the Analysis of New Frequently Purchased Consumer Products." *Operations Research* 18 (September–October):805–854.

Appendix 12A: Model Descriptions

Fourt and Woodlock (1960)

In this model, sales are projected by computing cumulative penetration and repeat ratios. The cumulative-penetration percentage is estimated through a simple growth curve that approaches a limiting value asymptotically. It is hypothesized that the increment in any period is a constant fraction of the difference between the ceiling value and the current level of penetration. To compute repeat-purchase rates, the authors compute a series of repeat ratios, which are defined as the number of households which purchased the product on the $(t + 1)$st occasion, having purchased it on the tth occasion. These repeat ratios are calculated for first repeat, second repeat, and so on. The repeat ratios are then successively applied to the estimated new buyers at the end of the observation period to arrive at an estimate of projected sales. The method of estimation of the growth curve is not described in much detail other than a footnote reference that the parameters are estimated by maximum-likelihood procedures.

Parfitt and Collins (1968)

The model's objective is to predict S, the long-run share of the new brand T in the test market

$$S = PRB$$

where P = ultimate penetration of brand

R = ultimate repeat-purchase rate for brand

B = buying-rate factor which models differences in purchase volume between buyers of new brand and buyers of product class in general; B is equal to 1.0 if buyers of the new brand buy it at the same rate as the rest of the product-class buyers[1]

To estimate P, the authors use a growth-rate model given by

$$K(t) = P(1 - e^{-at})$$

where $K(t)$ = cumulative penetration at time t

a = rate-of-growth parameter

Here P and a are estimated by discounted least-squares procedures. To estimate R, the authors seem to use subjective judgment rather than a statistical model.[2] And B is measured by the ratio of the amount of the product class bought by users of brand T to the amount of the product class bought by all buyers of the product class.

STEAM (Massy 1969)

The STEAM model consists of a primary model accompanied by a series of secondary models which describe the manner in which the parameters of the primary model vary across individuals, over time, and on the timing of purchases. The primary model specifies a probability function that governs the number of purchase events which take place by a certain time t (Poisson process). A secondary model is added to allow the mean purchase rate of the consumer to vary over time (nonhomogenous Poisson process). Next, the mean purchase rate is assumed to vary across individuals (the mean purchase rate is assumed to be distributed across individuals according to a gamma distribution). Finally, in order to account for the effects of the timing of purchases, the mean purchase rate is assumed to depend on the time of conversion to a particular depth of repeat. (Specificially, a parameter of the gamma distribution is made a function of the time of conversion.)

SPRINTER (Urban 1970)

In this model, consumers are classified into several "experience classes" as follows: potential trier, preference class, loyalty class 1, loyalty class 2, and nonloyal class. Within each class, the consumer passes through a five-step behavioral process: awareness, intent, search, choice, and postpurchase behavior. In the awareness step, consumers are classified into several states such as brand-aware, advertising-aware, product-appeals-aware, and so on. After being aware of the brand, consumer intentions are estimated for each awareness class. Search is determined by product availability. Choice is based on relative price and point-of-sale activity of the brand. Postpurchase behavior relates to issues such as word-of-mouth recommendations and forgetting. These five steps are assumed to be repeated over time, and depending on their choices, the consumers could move to different experience classes.

The model needs input from store audits; panel data on purchase history; survey data on awareness, preference, intent, use, word-of-mouth influence, salespersons' reports, media audit reports, and so on. The model's parameters are estimated by a variety of statistical procedures such as regression, nonlinear estimation, and a general conditional classification analysis program.

ESKIN (1973)

The model's objective is to project S_t, the cumulative sales of a brand at time t, and S_t is computed as follows:

$$S_t = \sum_{J=0}^{\infty} R_t(J)U_t(J)$$

where $R_t(J)$ = cumulative number of consumers repeating at least J times by period t

$U_t(J)$ = average units of purchased on Jth repurchase occasion

And $R_t(J)$ is calculated as follows:

$$R_t(J) = \sum_{i=1}^{t} \text{(cumulative fraction that repeats the } J\text{th time by period } t \text{ given that purchase } (J-1) \text{ was made during time period } i) \times \text{(number of consumers repeating the } J\text{th time in period } i)$$

The author posits that the use rate $U_t(J)$ is linearly related to the repeat-purchase level up to a ceiling value $\bar{U}$. Given the above relationships, the author hypothesizes certain behavioral relationships to estimate the cumulative fraction and the cumulative sales.

Thus, we see that the model ignores the transition from awareness to trial and models only the trial and repeat stages. However, Eskin's model analyzes consumers based on the time of their initial purchase and the number of times they have repurchased the new brand. In addition, the use rate is allowed to be different depending on the depth of repeat, that is, the number of times repeat purchases have been made.

Validation of the model is demonstrated by fitting it to a set of diary-panel data (the product class and the exact sample size are not given) available for 52 weeks. Estimation was done by using 12, 24, and 52 weeks

of data. For each data set, predictions were made for 12, 24, 36, and 52 weeks. Total sales, trial sales, repeat sales, trial transactions, and repeat transactions are forecasted and compared to actual data.

Nakanishi (1973)

In this model, the consumer is assumed to move through four stages in the adoption of a new product: pretrial, trial, acceptance, and rejection. Thus, with reference to figure 12–2, a fourth stage, namely the rejection stage, is added. The author postulates a transition probability matrix to describe a consumer's movement through the various stages. The waiting time in a stage (other than the rejection stage) is the same as interpurchase time and is treated as a random variable. A probability function is prescribed for interpurchase time (a nonhomogenous Poisson process), and the mean waiting time is made a function of marketing variables. The mean waiting time depends on the stage in which the consumer is in. Thus, the model allows the purchase frequency to be different for triers and repeaters. However, at the empirical stage the author assumes interpurchase time to be the same for the different stages.

Expenditures on network television, spot television, and direct-mail coupons are the three marketing-mix variables considered in the model. The effects of these variables are measured by fitting separate regression models for trial and repeat purchases.

NEWS (BBDO International, Inc., undated)

The new-product early warning system (NEWS) model developed by BBDO is an outgrowth of their earlier DEMON model (Learner 1968). The full details of NEWS have not been published. Consequently, we are limited to the sketchy information provided in the brochure that was available to us (BBDO International, Inc., undated) and the paper by Light and Pringle (1970). Unfortunately, the description of the model in Light and Pringle is not consistent with the description in the BBDO brochure. Since the brochure appears to be more recent, our description of NEWS is based on the description provided in the brochure.

The goal of the model is to forecast new-product sales and to provide diagnostic information to the manager. NEWS makes sequential forecasts of awareness, trial rate, and use rate. Awareness is linked to media weight, advertising impact, and promotional expenditures. Trial is related to awareness, product concept, promotion, distribution, and price. Repeat buying is based on product quality, distribution, and price. The sales forecast is made by estimating the use rate and using the trial and repeat-sales volume.

The data inputs to the model consist of a combination of panel data, subjective estimates, and other managerial inputs on price and media expenditures. Sensitivity analysis can be carried out to determine which parameters have a greater impact on the ultimate sales forecast. Finally, by integrating cost and other financial data, it is possible to evaluate the profitability of a particular marketing strategy. This implies that one should be able to determine optimal levels of the marketing-mix variables, and the NEWS brochure provides an illustration of the determination of the optimal advertising budget.

NEWPROD (Assmus 1975)

NEWPROD models all three stages of figure 12–2. A consumer can become aware of the new product through advertising, coupons, or free samples. An explicit model for describing the number of persons becoming aware through advertising is provided. The model allows for decreasing marginal returns to advertising. However, similar models are not provided for coupons or free samples. Depending on the method (that is, advertising, coupons, or free samples) by which they become aware, consumers who become triers are labeled as "purchase triers," "coupon triers," and "sample triers," respectively. A person who receives a free sample automatically becomes a trier. A trier becomes a repeater if he/she purchases the product within one week after the lapse of a purchase cycle from the time he/she became a trier. Consumers who become aware in time t but did not try the product in time $t + 1$ are classified as "reluctantly aware." Similarly, consumers who did not repeat are classified as "skeptical triers." Such groups of consumers have a lower probability of moving to the next stage in subsequent weeks.

The author is not very precise in providing modeling details and showing how the brand's market share is forecasted. Nor is he precise regarding data sources and estimation methods. He suggests that the different parameters could be subjectively estimated, obtained through test-market data or through pre-test-market data. He does mention regression analysis as a possible estimation procedure.

TRACKER (Blattberg and Golanty 1978)[3]

Model Structure

The model's goal is to forecast year-end sales for the new product three months after it has been introduced in the test market. The sales forecast is based on three models, one for each of the three stages of new-product acceptance: awareness, trial, and repeat.

The awareness model consists of the following relationship:

$$\ell n \frac{1 - A_t}{1 - A_{t-1}} = \alpha - \beta GRP_t$$

where α and β are parameters to be estimated, GRP_t is the gross rating points obtained by advertising, and A_t is the brand's total awareness in time period t. Note that the above relationship captures decreasing marginal returns to advertising.

In the trial model, two types of triers are considered: those who become aware in the current period and those who were aware in the previous periods but had not yet tried the product. Thus, the following model is proposed:

$$\Delta T_t = T_t - T_{t-1} = \alpha(A_t - A_{t-1}) + \beta(A_{t-1} - T_{t-1})$$

where α and β are parameters to be estimated, T_t is the *cumulative* percentage of triers of the new brand in period t, and A_t is the *cumulative* percentage of potential triers of the new product who are aware of the new product in period t. This model reflects the notion that newly aware potential triers have a higher probability of trial compared to those who were aware in the past but have not yet tried. In other words, the parameter α is expected to be larger than β (where α and β are both constrained to be between 0 and 1).

The above model of trial assumes that the new brand's price is the same as the average price in the product class. To allow for differential brand prices, the trial model is amended as follows:[4]

$$\Delta T_t^* = \Delta T_t / \bar{P}_t^{\gamma}$$

where ΔT_t^* = observed price-adjusted incremental trial

$\bar{P}_t$ = relative price (price of new brand divided by average price of product class)

γ = parameter to be estimated ($\gamma > 0$)

If the new brand's price is greater than the average industry price, then $\bar{P}_t > 1$ and $\Delta T_t^* < \Delta T_t$ (as expected). Similarly, $\Delta T_t^* > \Delta T_t$ if $\bar{P}_t < 1$ while $\Delta T_t^* = \Delta T_t$ if $\bar{P}_t = 1$.

Given incremental trial, trial-use rate (TU), repeat-use rate (RU), and

repeater proportions (r and $1 - d_i$), TRACKER computes sales for any period using the projection model. The projection model is illustrated in figure 12–3 for period 4. Total sales in period 4, $TS(4)$, could come from new triers in period 4 or from consumers who had tried the product in either period 1, 2, or 3 and who are repeating in period 4. The model estimates the percentage of consumers in each stage of repeat and computes each period's sales by aggregating the sales from each repeater class and the new trier class for that period. Year-end sales are obtained by summing the sales for each period.

Data Requirements

This model uses survey data which are collected by administering questionnaires to potential users of the product (sample sizes range between 500 and 1,000) in three waves sequenced every four weeks after the introduction of the new product (although this could vary depending on the average purchase cycle for the product). The nature of the information that is collected is indicated in table 12A–1.

Estimation

The model's parameters are estimated by pooling observations for different brands in a product category and assuming homogeneity of the parameters across the brands. The authors are forced to do this because they have only three observations per brand and estimating the parameters for the test brand would be impossible if only the data for the test brand were used. The awareness model's parameters are estimated by ordinary least squares, and the authors report R^2 values of 0.732 and 0.652 for two different product categories. In both product categories, the signs of the parameters match a priori expectations. The trial model is estimated by a nonlinear least-squares method. Again, the parameter estimates appear to be reasonable in terms of both signs and magnitudes. In the projection model, TU is set equal to 1 (because triers are likely to buy only one unit of the product), the $d(i)$ (see figure 12–3 for a definition) are estimated subjectively, while r and RU are estimated through the telephone surveys. The authors report that the subjective estimate of r in the first period agrees closely with the latter estimates. Therefore, an average value of r is used in the model. As far as the $d(i)$'s are concerned, a simulation conducted by the authors demonstrated that the predictions were sensitive to these estimates.

Table 12A-1
Information Available from the Questionnaire for the TRACKER Model

1. Product use	7. Other brands used
2. Frequency of product use	8. Frequency of purchase for each brand used
3. Unaided awareness of brand 1st mentioned 2nd mentioned Others	9. Rating of each brand
	10. Receipt of free sample
4. Aided awareness of brand	11. Use of coupon
5. Length of time since last purchase	12. Repurchase intentions 1st choice 2nd choice Others
6. Brand bought last	

Source: R. Blattberg and J. Golanty, "TRACKER: An Early Test-Market Forecasting and Diagnostic Model for New Product Planning," *Journal of Marketing Research* 15 (May 1978): 192–202.

Notes

1. In passing, we should mention a minor variation of the Parfitt-Collins model due to Ahl (1970). Ahl combines the P and B terms into a single term, the ultimate *volume* trial rate.

2. Shoemaker and Staelin (1976) do suggest a specific statistical model to estimate R.

3. TRACKER is described in considerable detail in this appendix. It appears to be the best of the nine models reviewed, and we felt that the reader would be interested in a more complete description.

4. The notation in the equation that follows is slightly different from the notation in Blattberg and Golanty (1978).

References

Ahl, David H. 1970. "New Product Forecasting Using Consumer Panels." *Journal of Marketing Research* 7(May):160–167.

Assmus, Gert. 1975. "NEWPROD: The Design and Implementation of a New Product Model." *Journal of Marketing* 39(January):16–23.

BBDO International, Inc. Undated. "How to Predict New-Product Success." Unpublished brochure.

Blattberg, Robert, and Golanty, John. 1978. "TRACKER: An Early Test-Market Forecasting and Diagnostic Model for New Product Planning." *Journal of Marketing Research* 15(May):192–202.

Eskin, Gerald J. 1973. "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model." *Journal of Marketing Research* 10(May): 115–129.

Fourt, Louis A., and Woodlock, Joseph W. 1960. "Early Prediction of Market Success for New Grocery Products." *Journal of Marketing* 25(October):31–38.

Learner, David B. 1968. "Profit Maximization through New Product Marketing Planning and Control." In *Applications of the Sciences in Marketing Management,* edited by Frank M. Bass, Charles W. King, and Edgar A. Pessemier. New York: Wiley, pp. 151–167.

Light, Lawrence, and Pringle, Lewis. 1970. "New Product Forecasting Using Recursive Regression." In *Research in Consumer Behavior,* edited by David T. Kollat, Roger D. Blackwell, and James F. Engel. New York: Holt, Rinehart and Winston, pp. 702–709.

Massy, William F. 1969. "Forecasting the Demand for New Convenience Products." *Journal of Marketing Research* 6(November):405–412.

Massy, William F.; Montgomery, David B.; and Morrison, Donald G. 1970. *Stochastic Models of Buying Behavior.* Cambridge, Mass.: M.I.T. Press.

Nakanishi, Masao. 1973. "Advertising and Promotion Effects on Consumer Response to New Products." *Journal of Marketing Research* 10(August):242–249.

Parfitt, J.H., and Collins, B.J.K. 1968. "Use of Consumer Panels for Brand Share Prediction." *Journal of Marketing Research* 5(May): 131–145.

Shoemaker, Robert, and Staelin, Richard. 1976. "The Effects of Sampling Variation on Sales Forecasts for New Consumer Products." *Journal of Marketing Research* 13(May):138–143.

Urban, Glen L. 1970. "Sprinter Mod III: A Model for the Analysis of New Frequently Purchased Consumer Products." *Operations Research* 18 (September–October):805–854.

13 Use of Consumer Panels for Brand-Share Prediction

J.H. Parfitt and *B.J.K. Collins*

After the launch or introduction of a new brand or special promotion of an existing brand, much can be gained from knowing the ultimate outcome long before it becomes apparent from observing historic data on sales figures or brand-share trends. When the outcome is apparent in the usual way, it is often too late for action that might have been effective earlier. There is great marketing value in an early warning prediction system. One such system is brand-share prediction using continuous data from consumer panels.

The objects of this article are:

1. to describe the method used to obtain a prediction of brand share from panel data.
2. to illustrate how much these predictions are likely to be valid in practice by drawing on case histories from several analyses.
3. to demonstrate some of the ways the prediction techniques have been improved by experience in the use of analyses.
4. to comment on some of the marketing lessons emerging from a study of these analyses.

The data used here are from analyses of the Attwood Consumer Panel results in Great Britain after 1960.

Basic Method of Predicting Brand-Share from Panel Data

The basic method was first described in some detail by Baum and Dennis [2]. Although several refinements have been made to improve the accuracy and speed the prediction date, the basic method is the same. In 1961, however, it was considered an analysis to be used only in studying newly-launched brands. It is now realized that the method can be used equally effectively to study existing brands.

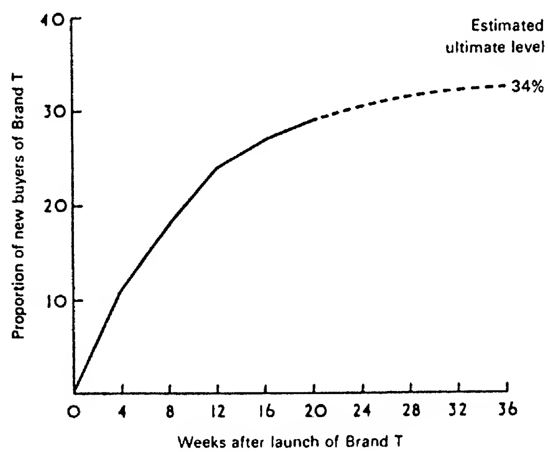
The raw data of the analysis are the continuous purchasing records of individual households. From these data, three basic components are selected:

Reprinted from J.H. Parfitt and B.J.K. Collins, "Use of Consumer Panels for Brand-Share Prediction," *Journal of Marketing Research* 5(May 1968):131–145, published by the American Marketing Association.

1. The cumulative growth in the number of new buyers of the brand or product being studied.
2. How often these new buyers buy the brand or product again after their first-recorded purchase. These repeat purchases of the studied brand are expressed as a proportion of the total purchases in the product field by these buyers; this is called the repeat-purchasing rate.
3. The rate of total product field purchasing of these particular buyers compared with the average of all buyers in the product field, i.e., are they usually heavy, light, or average buyers (by volume) in the product field. This is expressed as a buying level index, with the average buying level for the field being 1.0

Cumulative Penetration

To illustrate, assume a brand-share prediction analysis covers newly launched Brand T in an established and comparatively static product field (for example, toilet soap). The first requirement is to persuade households to try Brand T. If the advertiser cannot do this on a reasonable scale, there will be no future to predict. However, assuming reasonable retail distribution and a realistic advertising and promotion budget, getting housewives to try the brand is comparatively easy. The analysis, here, begins with the brand's launch. As each new Brand T buyer is picked up from panel records, she is recorded on a cumulative penetration curve similar to figure



Note: Total buyers of toilet soap = 100%.

Figure 13–1. Cumulative Penetration of Brand T

13-1. It is not necessary to wait until the total cumulative penetration has been completely observed before predictions can be made. Once the shape of the curve is determined a declining rate of increase is observed, it is possible to make a reasonable estimate of the ultimate likely penetration.

Repeat-Purchasing Rate

The ultimate success of any brand depends on the willingness of consumers, once having tried it, to continue purchasing it—normally this means to the exclusion (partial or whole) of the brands they were previously using. In oversimplied terms, persuading the consumer to try a brand is a function of distribution, advertising, and promotion, but getting her to keep using it is also a function of her acceptance of the product. The first is much more difficult to improve than the second. The new brand should be more than technically comparable with competitive brands already on the market. It must, for a significant number of consumers, be sufficiently more acceptable than the brands already used if they are to change their purchasing habits. The success or failure of the product's acceptance (against the background of advertising and promotion) is expressed by the repeat-purchasing rate.

As each new Brand T buyer appears on the panel records, her purchasing behavior in the product field is isolated as a continuous record over time. This is shown in a hypothetical buying pattern, table 13-1, in which Brands R and S represent the existing competitive brands.

The number of first-time buyers of Brand T accumulates to seven in the first three weeks of the study. The repeat-purchasing rate is calculated from the period after the first purchase of Brand T, here in two-week intervals. It is not an expression of calendar time because calculation begins for each buyer from his date of entry into the market, i.e., Buyer 2's first opportunity for repeat purchasing is in Weeks 2 and 3, but Buyer 7's first opportunity is in Weeks 4 and 5. In each of the two-week periods these seven buyers made ten purchases in the product field of which successively 60 percent, 50 percent, and 40 percent were repeat purchases of Brand T. This pattern of a declining repeat-purchasing rate is normal because second, third, fourth, etc., purchases of a new brand still tend to be exploratory and are also often still a benefit from promotional activity of the launch. The critical point for the prediction of the ultimate brand share comes when the repeat-purchasing rate begins to level off, as in Figure 13-2. At this point, considering an estimate of the ultimately likely penetration, it is possible to calculate what the brand share will eventually be.

Table 13–1
Illustration of Repeat-Purchasing Rate Calculation from a Hypothetical Example

Weeks:	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
Buyer:											
1	*T*	*T*	*R*	*T*	*R*	*R*	*R*	*R*	*T*	*R*	*R*
2	*T*		*S*		*S*		*S*		*S*		*S*
3		*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*
4		*T*		*T*		*R*		*T*		*R*	
5			*T*		*T*		*T*		*T*		*T*
6			*T*	*T*	*R*	*T*	*R*	*R*	*S*	*R*	*R*
7			*T*	*S*		*S*		*S*		*S*	
Cumulative Buyers	2	4	7								
Repeat-Purchasing Rate				6/10 = 60%		5/10 = 50%		4/10 = 40%		4/10 = 40%	

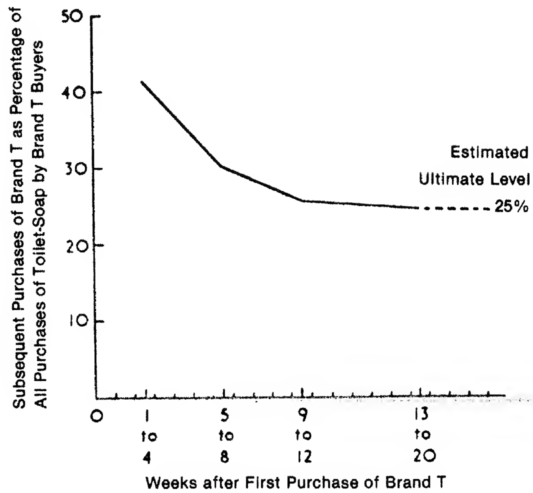


Figure 13–2. Repeat-Purchasing Rate for Brand T

Buying Level Index and Prediction Calculation

It is estimated in figure 13–1 that Brand T is likely to reach approximately 34 percent of potential buyers (a potential buyer here is a buyer of toilet soaps). In figure 13–2 the repeat-purchasing rate for these buyers has leveled off near the 25-percent level. With the assumption that households having tried Brand T purchase toilet soap on an average at the same rate as the average of total buyers in the market, then the predicted share for Brand T is:

estimated penetration × repeat-purchasing rate × buying rate index. . . . (i)

in this example this is $34\% \times 25\% \times 1.00 = 8.5\%$.

The buying level of buyers of Brand T could be an important element in the calculation because if the households who tried Brand T were particularly heavy buyers of toilet soaps (maybe an index of 1.2 compared with the average of all toilet soap buyers), or were particularly light buyers (maybe an index of 0.8) then this would considerably influence the calculation of the ultimate share of Brand T. This is illustrated as follows:
Heavy buyers of toilet soap:

$$34\% \times 25\% \times 1.20 = 10.2\%$$

Average buyers of toilet soap:

$$34\% \times 25\% \times 1.00 = 8.5\%$$

Light buyers of toilet soap:

$$34\% \times 25\% \times 0.80 = 6.8\%$$

Assumptions Underlying the Prediction

Two basic assumptions are implicit in the prediction calculation:

1. The retail distribution of the new brand is uniformly high in the area under study or, failing that, it is not substantially worse now than it is likely to be in the foreseeable future.
2. Besides the advertising and promotional activity accompanying and during the brand's launch (including competitors' retaliatory measures), the circumstances of the market will remain much the same in future as they have been during the prediction measurement.

This, of course, is asking a lot of any market in reality. Actually, observations of the relation between predicted shares and actual shares achieved over periods of up to two years suggest that the prediction will be a reliable guide to likely ultimate share in almost all cases in markets where no major changes take place. If a major change occurs in the market after the prediction has been made—such as the launch of another new brand that is successful—then the prediction will no longer be valid because the new circumstances of the market will not have been considered in the calculation.

How Valid is Brand-Share Prediction?

Two questions inevitably arise from the use of brand-share prediction analyses:

1. Are the predictions correct, i.e., does the brand really settle down around the share predicted for it?
2. Even if the predictions are correct, can the prediction calculation be made before it is clear from other data what will happen to the brand anyway, i.e., can the prediction be made in sufficient time to suggest a course of action that might not be otherwise obvious from other available data such as sales figures and current brand-share measurements?

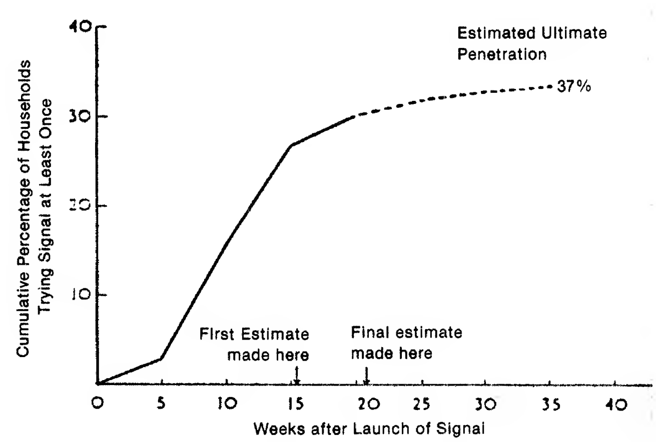
These two questions are studied later with actual case histories.

How does brand-share prediction work in practice? Look at an actual prediction calculation made by Baum and Dennis [2] to see how the brand progressed after the prediction had been made. The penetration and repeat-purchase patterns of the brand (which is Signal toothpaste) are shown in figures 13-3 and 13-4. The predicted share for Signal was 40 percent of 37 percent = 14.8 percent, assuming the total field-buying rate for Signal buyers was about average. Unfortunately, since the buying level index refinement was introduced sometime after this analysis, the buying level index for this case is not known.

Now look at these: (a) at what stage in the market development for Signal the prediction was made, and (b) at what level the Signal share found equilibrium, and for how long it remained there.

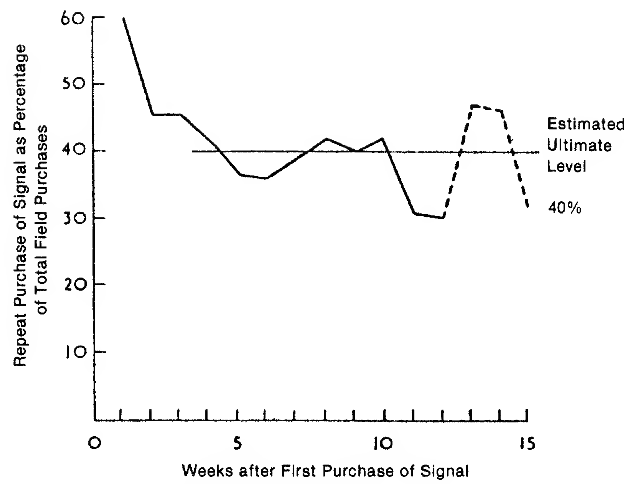
The first point is best studied by looking at four-week shares for Signal during the launch period (figure 13-5). Here the position is slightly complicated since the launch was phased across the country over a 12-week period which somewhat delays the point of final prediction.

Thus, the prediction was made as a first estimate as the brand reached its peak, and as a final estimate as it began its inevitable decline from the peak, i.e., between 16 and 20 weeks after launch. (The length of time required, after the start of the analysis, for a prediction to be made depends on two factors: (a) the average frequency of purchase in the product field,



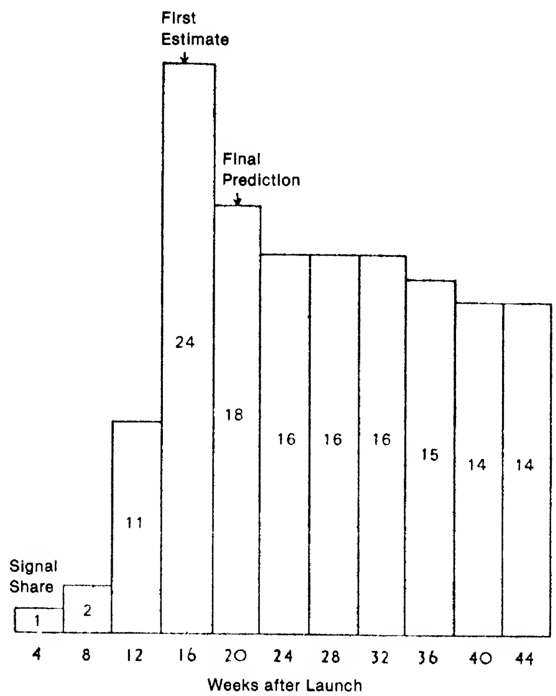
Source: Attwood Consumer Panel, Great Britain.
Note: Total dentifrice buyers in 26 weeks = 100%.

Figure 13-3. Cumulative Penetration of Signal



Source: Attwood Consumer Panel, Great Britain

Figure 13–4. Repeat-Purchasing Rate of Signal Buyers



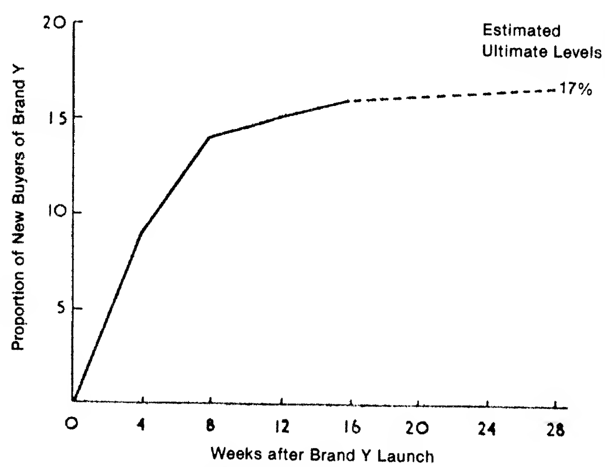
Source: Attwood Consumer Panel, Great Britain.

Figure 13–5. Signal's Share of Total Dentrifrice Market in the First 11 Four-Week Periods after Launch

and (b) the concentration of advertising and promotion used to obtain rapid market penetration by the brand.) The predicted level was reached between 36 and 40 weeks after launch; the product stayed at this general level (between 14 and 15 percent) for two years following the prediction. The launch of the fluoride toothpastes in the middle of Year 3 upset the market's equilibrium sufficiently to destroy the basis on which the Signal share had been predicted. Thus the prediction of the Signal share was made just after the peak of the launch but long before it could be determined where the share level would stabilize. The prediction remained valid for two years in a highly competitive market. To what extent is this a reliable guide to the use of these prediction analyses, or was it accidental?

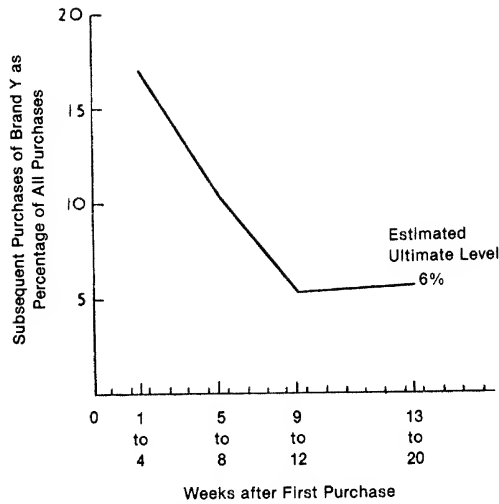
Before study of the evidence, it should be recognized that a substantial proportion of brands do not reach satisfactory share levels and are withdrawn. In these cases the brand-share prediction analysis indicates the low ultimate share, but the accuracy of the prediction cannot be proved because the brand does not stay on the market. So, prediction analyses first indicate the likelihood of success or failure, and only in adjudged successes does the opportunity usually occur to assess the accuracy of the prediction. Failure generally takes the form of exceptionally low repeat-purchasing rates, i.e., the brand makes a reasonable penetration into the market but very few people continue to purchase it. Figure 13-6 and 13-7 give actual examples from a toiletry product.

Thus, a penetration of 17 percent of the market, though not all that healthy, could nonetheless form the basis for a viable brand share. If, for



Source: Attwood Consumer Panel, Great Britain.
Note: Total buyers in the product field = 100%.

Figure 13-6. Cumulative Penetration of Brand Y



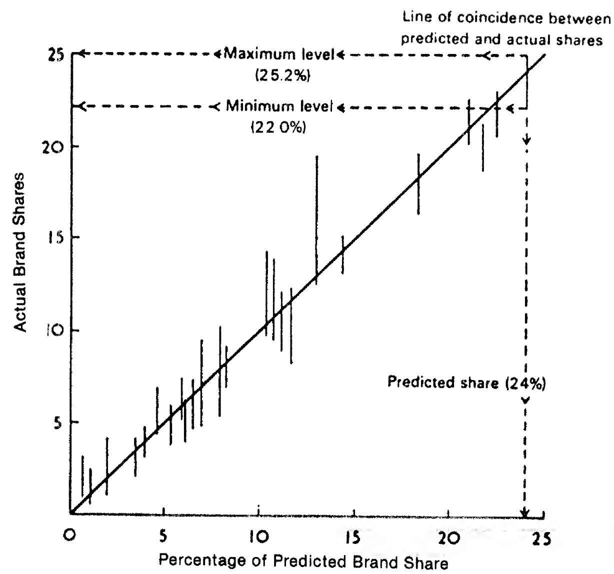
Source: Attwood Consumer Panel, Great Britain.

Figure 13–7. Repeat-Purchasing Rate of Brand Y Buyers

instance, the repeat-purchasing rate had been 35 percent (less than that for Signal) on a penetration of 17 percent, this would produce a predicted share of around 6 percent. In this particular field, already well-fragmented with competitive brands, this could be regarded as a reasonably successful launch. In fact, the repeat-purchasing rate stabilized at the very low position of 6 percent, and 6 percent of 17 percent produces a predicted share of 1 percent—far too low to be viable for a nationally promoted brand in this field (and too low for need to apply the buying level index). Clearly the fault lies with the repeat-purchasing rate; buyers having tried the brand did not feel disposed to use it again. In other words, the product did not gain acceptance.

So, prediction of failure has to be taken at its face value, but prediction of success can be studied further on accuracy—by studying what actually happened to the brand share after the launch. The results of some 24 successful ones where share predictions were made and the field was measured on the Attwood panel for the following 18 months are shown in figure 13–8.

The critical period, representing that for which the share prediction was to apply, is the six-month period between 12 and 18 months after the launch. Thus the predicted brand share (single-valued) is compared with the range of actual brand shares achieved (based on 6 × 4 weekly periods). The diagonal line on the chart passes through all points where predicted and actual brand shares coincide. The vertical bars show the range of actual four-week brand shares achieved between 12 and 18 months after the launch. Thus the extreme top right-hand example on the chart had a



Source: Attwood Consumer Panel, Great Britain.
Note: Based on six four week reports one year later.

Figure 13-8. Comparison between Predicted and Actual Brand Shares

predicted share of 24.0 percent (immediately afterward) and achieved actual shares in the period 12–18 months afterward ranging from a maximum of 25.2 percent to a minimum of 22.0 percent. The fields represented in this summary of prediction analyses cover products like dishwashing liquid, toilet tissue, butter, instant coffee, and floor polish, and include predictions in static total markets and in very dynamic total markets.

Of course, actual brand shares are seldom completely static—particularly in product fields with considerable promotional activity and particularly on a measurement as frequent as a four-week one (shares that can be quite static on an annual average basis will often show considerable fluctuations on a four-week basis). The share prediction, therefore, although it is expressed as a single value, really means the brand share will eventually assume this level as an average. All estimates of ultimate level of brand share are subject to the necessary condition that the market in which the new product is launched remain relatively stable in the period after the product's introduction. Figure 13-8 shows these predictions can be extremely accurate, particularly in relatively stable markets.

Some Refinements of the Technique

Earlier Prediction of Estimated Ultimate Level of Penetration

It is possible to use a marketing model of penetration that allows an early prediction of the expected ultimate level. Details of the model's statistical basis and an indication of the method of fitting the model to the actual data and the accuracy obtained, age given in appendix 13A.

The model's application enables an estimate of ultimate penetration to be made before the curve shows any marked tendency to level off. This, in itself, is valuable in planning marketing strategy. It is also helpful in reducing the length of the minimum study period necessary to produce a first estimate of expected brand share. The selected minimum time period is thus dependent on the validity of the secondary marketing variable, the repeat-purchasing rate. This measure is dependent on the frequency of purchase of the field containing the launched brand and may thus be determined before the study begins, assuming a reasonably smooth launch for the new brand.

In many instances the penetration curve has already begun to level off at the end of the selected minimum time period even though it has then been found that the deduced ultimate level based on the fitted model may be at variance to that produced by free-hand extrapolation. Some examples of actual penetration curves are given in appendix 13A to illustrate this point. However, the greatest advantage in using the penetration model is that the projected estimates provide an expected growth of new buyers with which it is possible to compare the subsequent actual penetration. The differences may then be attributed to such marketing disturbances (extensive price-cutting, free sampling, couponing, increased advertising) as are evident, in the form of a quantitative measure of additional penetrations and, after further analyses, additional brand share.

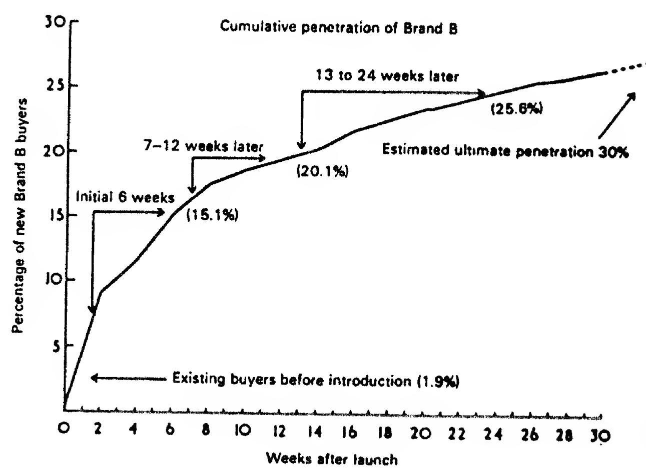
Allow for the Launch of a New Product/Field

Occasionally a new product cannot be considered a constituent part of an established field. Then, it is obviously not possible to apply the brand-share technique. However, the repeat-purchasing rate can be redefined in terms of quantity or expenditure per four-week period for a buying household and thus estimate an ultimate sales level. This calculation is, of course, possible in all prediction analyses, but it serves no advantage over estimations by brand share and, therefore, is only used when the product under study has no clearly defined market on which to base share.

Repeat-Purchasing Rates at Different Stages of Entry

A prolonged study of repeat-purchasing rate behavior has led to the conclusion that on average the sooner a buyer enters the market for a particular brand, the higher will be that buyer's repeat-purchasing rate. That is, the early entrants to a new brand's market tend to have disproportionate importance in contributing to the ultimate share that brand will achieve. For example, Brand B, an edible fat brand, is launched in one commercial TV area. It was already available in the area before its official launch because the distribution to other areas overlapped into it. Six months after launch, its estimated penetration in the market was 30 percent, the repeat-purchasing rate had leveled at 15.5 percent, and the buying rate index of Brand B buyers was 1.05, producing an estimated brand share of 4.9 percent. When this analysis is done separately by the buyers' date of entry into the market, the ultimate share prediction has sharply contrasting elements (see figure 13–9 and table 13–2). This conclusion has important marketing connotations to be considered later.

Thus, the longer after the launch date a buyer entered the market for Brand B for the first time, the lower the average repeat-purchasing rate and the smaller the relative contribution to the ultimate brand share. This pat-



Source: Attwood Consumer Panel, Great Britain.
Note: Total buyers in one product field = 100%.

Figure 13–9. Cumulative Penetration of Brand B Analyzed by Buyers' Date of Entry into Market for Brand B

Table 13-2
Repeat-Purchasing Rate for Brand B Analyzed by Date of First Entry into Market

Buying Group by Date of Entry	*Contribution to Cumulative Penetration of Brand B by Each Group of Buyers*	*Repeat-Purchasing Rate of Each Buying Group*	*Buying Rate Index*	*Contribution of Each Group to Estimated Brand Share*
1. Existing buyers before Brand B was officially launched	1.9%	40%	1.21	0.9%
2. First-time buyers in the first 6 weeks	13.2	17	1.00	2.2
3. First-time buyers in the second 6 weeks	5.0	15.5	1.12	0.9
4. First-time buyers 13–24 week period	5.5	8.5	1.08	0.5
5. Estimated first-time buyers after 24 weeks	(4.4)	(8.5)	(1.00)	0.4
Total (or average)	30.0%	15.5%	1.05	4.9%

Source: Attwood Consumer Panel, Great Britain.

tern is consistent in all these analyses, and analyzing the repeat-purchasing rate by the point of entry on the cumulative penetration curve is now standard practice. It has three considerable advantages:

1. Greater accuracy is obtained in the ultimate brand-share prediction calculation because the repeat-purchasing rate of new buyers estimated still to enter the market can be determined at the latest marginal rate rather than the average rate (that would tend to overestimate).
2. Knowledge of the repeat-purchasing rate of the most recent entrants to the market can indicate whether there is any point in trying to make special efforts to increase the brand's cumulative penetration.
3. It contributes much to the understanding of factors that determine the brand's share and to marketing activities aimed at changing that share.

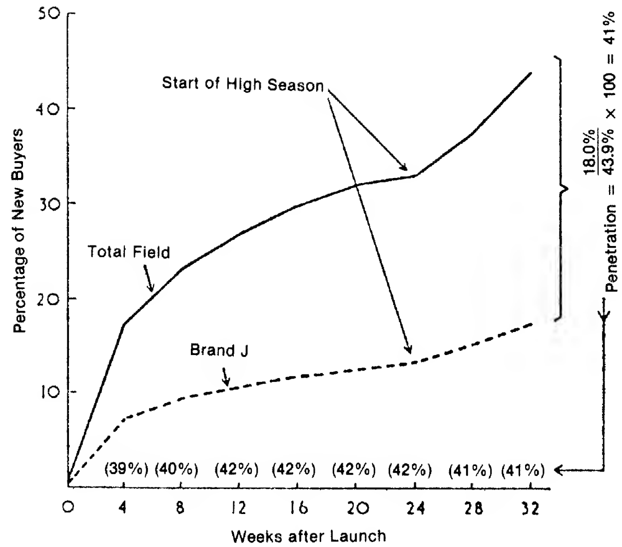
Total Market Cumulative Buying Levels to Determine Share Predictions in Changing Markets

An assumption not yet mentioned in relation to brand-share prediction calculations is that the total market level is not affected by the launch of the new brand. For a wide range of product fields that assumption is substantially true, but for some it is not, and in these cases, it could seriously affect the accuracy of a prediction calculation if it is not considered. Highly seasonal markets are special and require the same analytical techniques for correction.

To allow for these possibilities, the accumulation of both total buyers of the product field and buyers of the studied brand is produced. It is only when the rate of increase in accumulation of buyers for the brand exceeds the corresponding rate for the total market that the brand's marginal penetration is contributing anything to its ultimate level of penetration. In figure 13–10, although the penetration of the brand appears to surge suddenly upward, it is only the result of seasonal factors. The true penetration does not alter at all.

Cumulation of Case History Data Enables Quicker Assessment

One result of many brand-prediction analyses in the last six years is that something approaching a set of rules begins to emerge from the data. From these it is possible to determine relatively early whether the pattern of the brand being measured shows elements of success or failure. They are not rigid rules but have considerable value in two ways:



Source: Attwood Consumer Panel, Great Britain.
Note: All households = 100%.

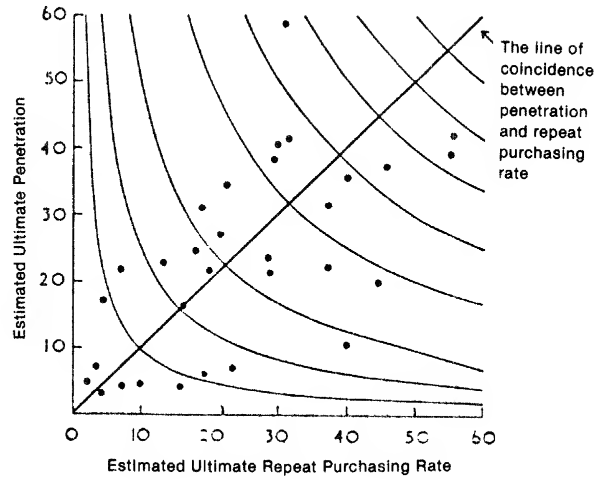
Figure 13-10. Isolation of Seasonal Factors in Cumulative Penetration of a Brand (Brand J)

1. A relatively small cumulative penetration accompanied by a relatively high repeat-purchasing rate might suggest that special promotional activity or improved distribution could make a substantial improvement in ultimate brand share.
2. The exact relationship between a relatively high accumulation of buyers and a low repeat-purchasing rate can help explain whether a product deficiency or a more flexibe factor is at fault.

One thing is certain—there is no rule about the level of repeat purchasing to be expected at different levels of penetration. There is no natural equilibrium level of repeat purchasing associated with particular levels of penetration, shown in figure 13-11 where several examples of the relationship of repeat-purchasing rates to penetration levels are plotted, with contour lines of constant brand share.

Applying Brand-Share Prediction Analysis to Established Brands

It was assumed, when these analyses were first devised, that they would apply only to measuring progress of a new brand from its launch; for a long



Source: Attwood Consumer Panel, Great Britain.
[a]Multiplied by the buying rate factor.

Figure 13–11. Relationship of Repeat-Purchasing Rates to Penetration Levels in 31 Case Studies

time the analyses were used only for this purpose, probably for two reasons:

1. Conditions existing when the brand was launched for new buyer accumulation and repeat purchasing are somewhat different from those applying to an established brand.
2. Even if the above is not true, established brands do not have the dramatic changes accompanying the launch of a new brand, and there would, therefore, be nothing for a share-prediction analysis to do.

It has since been observed that movements below the surface of an established brand differ only in degree rather than kind from those present at the launch of a new brand. Share-prediction analyses have been of considerable use, for example, accompanying major promotions of existing brands to determine how much ultimate effect they are going to have on a brand share. Some examples are shown later. They are now becoming important in the measurement of loyalty patterns for established brands—since, of course, the repeat-purchasing rate is a way of expressing loyalty. The only technical difference in analysis for an established brand, compared with a new brand, is that it begins at a fairly arbitrary point in time (although its selection is usually related to the start of a promotion) and the accumulation of buyers is based only on buyers picked up after that time. From then it is exactly like an analysis for a new brand, even to the point

when the repeat-purchasing rate is highest for buyers who come in first (here, of course, buyers who come in first will include a high proportion of regular and loyal buyers).

Marketing Factors

The purpose of test marketing is to try to determine the likely success or failure of a product before it is committed to a national launch. Using a brand-share prediction analysis in a test marketing situation will provide certain advantages:

1. It will determine within basic preset limits whether the brand is likely to succeed or fail.
2. If success is predicted, it will determine, quite accurately, what degree of success expressed in brand share terms is likely to be achieved. This has obvious advantages in production, advertising, and promotion planning.
3. The prediction will be made long before the same conclusions could be drawn from sales or standard research data. Since timing is a considerable factor in the value of a test marketing operation, early prediction of results is clearly important.

However, certain marketing conclusions emerging from a study of these analyses go beyond the basic conclusions of the analysis itself. That is, the later a buyer enters the market for the studied brand, the lower her repeat-purchasing rate is likely to be. The importance of this conclusion to marketing is considerable, and the range of its application must be exploited. To illustrate this point, let us make an unreasonable statement—at least the statement is unreasonable in its simple form—although it is often an implicit assumption in certain marketing decisions made by marketing men who would recognize its unreasonableness in the simple form: "My brand has achieved a 20 percent penetration of the market, with a 30 percent repeat-purchasing rate. This will give me an ultimate share of 6 percent of the market. If I could double my penetration of the market, I would push my brand share up to 12 percent."

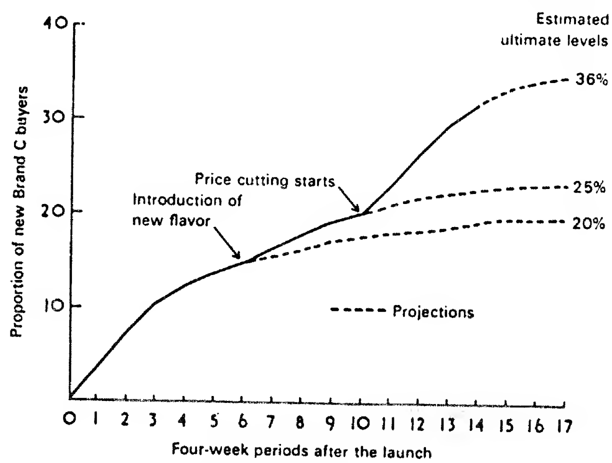
The implicit assumption, is that an increase in penetration can be achieved without loss in repeat-purchasing rate. The only possible exception follows. If the repeat-purchasing rate were kept artificially low by a blockage in distribution, and if the distribution were then substantially improved, then the repeat-purchasing rate might remain constant when the penetration doubled. In fact, assuming distribution remains fairly constant, the effect of doubling the penetration in the example could be anything from hardly any increase in ultimate share (maybe from 6 percent to 6.5

percent) to a substantial increase (unlikely to exceed a movement from 6 percent to 9 percent). Which result it is seems to depend considerably on the circumstances of the launch to date and the methods used to improve the cumulative penetration.

A more sensible statement is "it is comparatively easy within limits to influence cumulative penetration, i.e., the number of first-time buyers, but it is extremely difficult to create or influence repeat purchasing for any length of time." The time for influencing repeat purchasing is at the product development stage. By the time it reaches the market, there is only very limited room to maneuver, particularly if the product has comparatively low acceptance among new consumers compared with other competitive products. Some examples follow.

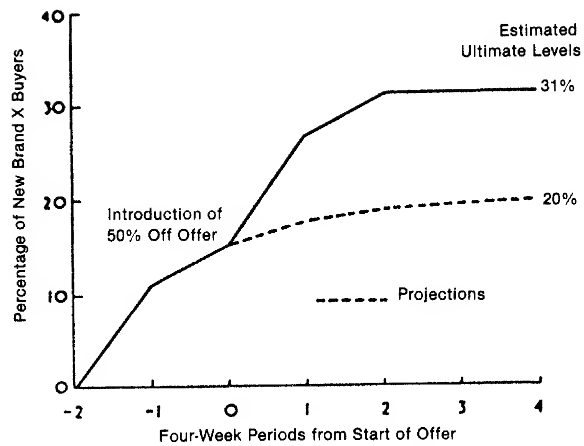
Deep Price Cutting as a Means to Improve Penetration: (a) a new brand, Brand C, using attractive retailer-trade terms to reduce the consumer price, at a time when the penetration curve was leveling off, and (b) an established brand, Brand X, using a 50 percent price-cut offer, also when the penetration curve was leveling off. In each case there is a sharp increase in penetration associated with price cutting, running against the normal direction of a penetration curve (shown in figures 13-12 and 13-13). The net increases in penetration associated with the offer were: Brand C—11% (43% increase), Brand X—11% (55% increase).

The repeat-purchasing rates of buyers brought in before, and after, the price cuts, respectively, appear in table 13-3. In both cases the price cutting



Source: Attwood Consumer Panel, Great Britain.
Note: Total buyers in the product field = 100%.

Figure 13-12. Influence of an Offer to Retailers on Cumulative Penetration of Food Brand C



Source: Attwood Consumer Panel. Great Britain.
Note: Total buyers in the product field = 100%.

Figure 13-13. Influence of a 50 Percent Price Cut on Cumulative Penetration of Detergent Brand X

considerably increased the number of first-time buyers but added proportionately much less to the ultimate brand share. The reason is that new buyers gained thus showed relatively little inclination to repeat purchase. In effect, it appears that they were only prepared to be in the market at the substantially reduced price. When the price moved back towards normal, they dropped out. Particularly with Brand X this is a clear, if extreme, example of the ease with which penetration can be influenced but not repeat purchasing.

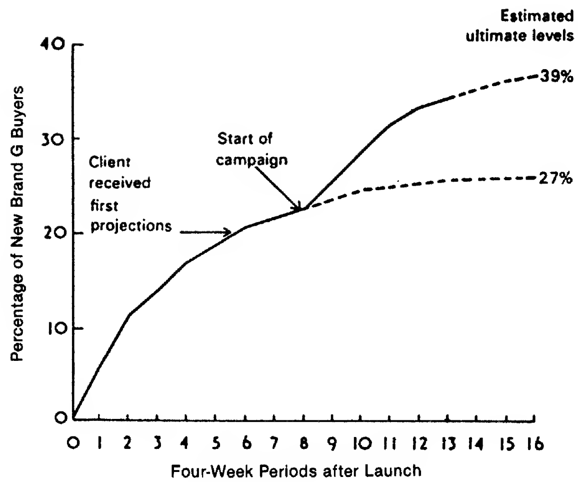
Other Promotions to Improve Penetration: (a) a new brand, Brand G, and (b) an established brand, Brand H. In both, special promotions were used when the penetration curve for the brand was leveling off, and in both, the promotions produced a substantial and unseasonal increase in penetration (see figures 13-14 and 13-15). The net increases in penetration associated with the promotions were: Brand G—12% (44% increase), Brand H—35% (117% increase).

The repeat-purchasing rates of buyers brought in before and after the promotions, respectively, are shown in table 13-4. In these examples, the promotions had been successful in adding buyers, but to a greater extent than the price-cutting examples; they had also been successful in achieving a healthy repeat-purchasing rate among these new buyers, and therefore, made significant increases in the ultimate brand share.

Table 13–3
Estimates of Ultimate Market Shares Likely to be Achieved by Brands C and X, Both with and without Price Cutting

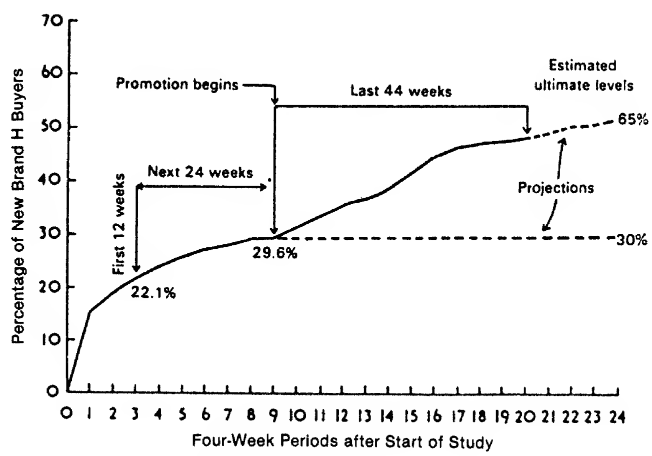
Category	*Ultimate Penetration* ×	*Repeat Purchasing* =	*Ultimate Brand Share*	*Buying/Rate Index*
		Brand C[a]		
Before introduction of price cutting				
Before new flavor	15%	20%	3.0%	
After new flavor	5	18	0.9	
Expected further buyers	5	18	0.9	
Total	25	19	4.8	
After introduction of price cutting				
Before offer	20	19.5	3.9	
After offer	13	10	1.3	
Expected further buyers	3	10	0.3	
Total	36	15.3	5.5	
		Brand X[b]		
Before offer's introduction	20	25	5.15	1.03
After offer's introduction				
Buyers without offer	20	25	5.15	1.03
Buyers with offer	11	6	0.65	1.01
			5.8	

Source: Attwood Consumer Panel, Great Britain.
[a]Apparent effect of offer on ultimate brand share = 0.7% (or 15% increase).
[b]Apparent effect of offer on ultimate brand share = 0.65% (or 12½% increase).
[c]Share calculation allowing for buying-rate index.



Source: Attwood Consumer Panel. Great Britain.
Note: Total buyers in the product field = 100%.

Figure 13–14. Influence of Promotion on Penetration of Brand G



Source: Attwood Consumer Panel, Great Britain.
Note: Total buyers in product field = 100%.

Figure 13–15. Influence of Promotional Activity on Penetration of Established Brand H

The purpose of these two sets of examples is not to demonstrate that substantial price cutting will do very little to ultimate brand share, and other promotional activity will be more successful. This may be true, particularly in product fields with a high elasticity of demand related to price,

Table 13–4
Estimates of Likely Ultimate Market Shares by Brands G and H, with and without Promotional Activity

Category	*Ultimate Penetration*	× *Repeat Purchasing*	= *Ultimate Brand Share*	
		Brand G[a]		
Before promotion				
Buyers already in	24%	34%	8.2%	
Expected further buyers	3	25	0.75	
Total	27	33	8.95	
After promotion				
Buyers already in	24	34	8.2	
Buyers coming in after the promotion	15	17	2.6	
			10.8	
	Ultimate Penetration	× *Repeat Purchasing*	− *Buying-Rate Factor*	= *Ultimate Brand Share*
		Brand H[b]		
Before the promotion				
First 12 weeks	22.1%	50%	1.29	14.3%
Next 24 weeks and expected further buyers	7.9	29	0.85	2.0
Total	30.0	46	1.18	16.3
After the promotion				
First 12 weeks	22.1	50	1.29	14.3
Next 24 weeks	7.5	29	0.85	1.8
Last 44 weeks and expected further buyers	35.4	20	0.90	6.4
Total	65.0	34	1.03	22.5

Source: Attwood Consumer Panel, Great Britain.

[a]The promotions' apparent effect on the ultimate brand share = 1.85% (or 23% increase).

[b]The promotions' apparent effect on the ultimate brand share = 6.2% (or 38% increase)

but there are insufficient examples here to prove it. What the examples do show is that it is much easier to increase penetration than to improve ultimate brand shares and that the extent to improve ultimate brand shares and that the extent which a brand-share improvement follows from an increase in penetration appears to depend on the methods used to increase penetration. The examples also demonstrate that brand-share prediction can be very effective in analyzing purchasing patterns for established brands as well as new brands.

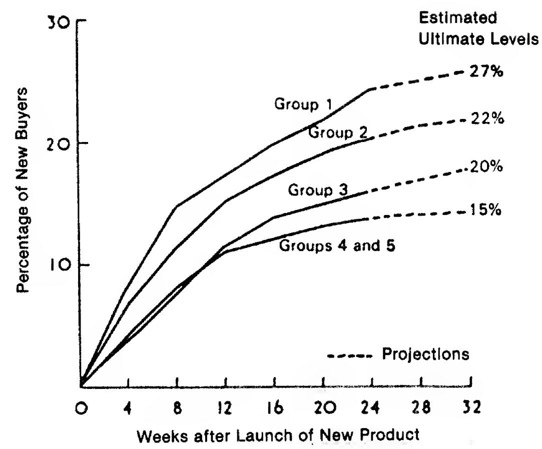
Postscript

Psychological Grouping Analyses for Brand-Share Prediction

The Attwood Consumer Panel housewives are classified by several psychological or attitudinal categories based on attitude statements rated by the housewives on a five-point scale. From her reactions to the statements in a particular attitude area each housewife was given a position on that scale. For this analysis purpose the scale "willingness to experiment in shopping" has been used. The penetration of a new product and the subsequent repeat-purchasing rate have been analyzed separately for each of four groups on the scale, ranging from those classified as "most willing to experiment" to those "least willing to experiment." The proportion of total housewives falling into each group follows:

Group	*Rating*	*Percent*
1	Most willing to experiment in shopping	18
2		30
3		37
4 & 5	Least willing to experiment in shopping	15

The penetration of the new product into each group on this scale is shown in figure 13-16. Thus, a higher proportion of Group 1, "most willing to experiment in shopping," was prepared to try the new product than any of the other groups. The proportion diminishes steadily towards Groups 4 and 5, "least willing to experiment in shopping," in which the proportion trying the new product is only a little over one-half of that in Group 1. Since this is a new product, with an ill-defined total market, the repeat-purchasing rate is expressed in absolute quantity terms, but this makes no difference to



Source: Attwood Consumer Panels, Great Britain.
Note: Total households in specified group = 100%.

Figure 13-16. Cumulative Penetration of a New Product/Field by "Willing to Experiment" Attitude Groups

the deduced conclusions. For each of the four groups, the estimated ultimate rate of repeat purchasing follows:

Group	*Rating*	*Repeat-Purchasing Rate*[a]
1	Most willing to experiment in shopping	28
2		35
3		26
4 & 5	Least willing to experiment in shopping	10

[a]Given in terms of packages per four-week period per 100 buying households.

Thus, although Group 1 members were more prepared to try the new product the first time, they were relatively less willing to continue buying the product than Group 2 housewives. Groups 4 and 5 were not only the least willing to try the new product the first time, they were also the least willing to continue buying it. Groups 1 and 2 made the biggest relative contributions in terms of the ultimate level of product purchases (see table 13-5). Although Groups 1 and 2 make a similar relative contribution to the ultimate level of purchases, i.e., they contribute some 30 percent more to the ultimate purchasing level than their proportion in the population, they

Table 13–5
Relative Contributions Made to Ultimate Buying Level of New Product by Housewives Classified by Their Willingness to Experiment in Shopping

Group	*Total House-wives (%)*	*Penetration*		× *Repeat Purchasing*	= *Ultimate Buying Level*[a]	
1. "Most willing to experiment"	18	(27%)	4.9%	28	1.4	24%
2.	30	(22)	6.5	35	2.3	39
3.	37	(20)	7.4	26	1.9	33
4 & 5. "Least willing to experiment"	15	(15)	2.3	10	0.2	4
	100	(21%)	21.1%	27.5	5.8	100%

Source: Attwood Consumer Panels, Great Britain.
[a]Expressed in packages per four-week period per 100 households.

do so by different paths. Group 1 housewives show a greater tendency to try the product in the first place but less tendency to continue buying it compared with Group 2 housewives, who are more cautious about trying it, but once having done so, are more inclined to continue using it.

It is possible that a heavy concentration of advertising and promotion on housewives in Groups 1 and 2 (the half of the population with above-average willingness to experiment in shopping) would increase the ultimate buying level of the product. However, many more analyses of this kind are needed before these abstract concepts can be used with any confidence. At this stage they serve only as direction into an area of isolating new brand triers who continue to use the product.

Conclusions

This article attempted to explain five main points:

1. Brand-share prediction, derived from continuous consumer panel data, works. Many analyses conducted from Attwood panel data confirm a close relationship between predicted and observed shares.
2. Prediction, itself, can be made long before the stabilization of share and, hence, is an exceptionally valuable marketing tool.
3. Refinements of the prediction technique introduced since the first studies in 1960 have improved the accuracy of prediction, advanced the prediction date and increased the understanding of the underlying characteristics of consumer purchasing behavior.
4. Analyses over a wide range of product fields suggest the existence of several rules of consumer behavior important in aiding marketing decisions.
5. There is still much to learn from the ramifications of this basically simple technique, particularly in attitude groupings, loyalty studies, and the relationship of advertising and promotion to ultimate brand share.

References

1. F.J. Anscombe, "Estimating a Mixed-Exponential Response Law," *Journal of the American Statistical Association,* 56 (September 1961), 493–502.
2. J.Baum and K.E.R. Dennis, "The Estimation of the Expected Brand Share of a New Product," ESOMAR Congress 1961.
3. Louis A. Fourt and Joseph W. Woodlock, "Early Prediction of Mar-

ket Success for New Grocery Products," *Journal of Marketing,* 25 (October 1960), 31–8.

4. W.G. Gilchrist, "Methods of Estimation Involving Discounting," *Journal of the Royal Statistical Society,* Series B, 29, No. 2, (1967), 355–69.

Appendix 13A

Objective

This appendix seeks to:

1. formally define terms and visually explain the underlying brand-share prediction model.
2. briefly describe the statistical method used for forecasting the penetration.
3. indicate the degree of predictive accuracy achieved.

Brand-Share Prediction Model

Defining t as a discrete time variable, measured after the launch of the new product (Brand α) within a defined product field (Field A), then:

$N(t)$ is the number of new buyers of Brand α introduced at time t,

$F(t)$ is the number of new buyers of Field A introduced at time t,

$P(t)$ is the penetration of Brand α within Field A at time t,

$$= \frac{\sum_{r=0}^{t} N(r)}{\sum_{r=0}^{t} F(r)} \times 100,$$

(i.e., as a percentage).

P is the ultimate penetration of Brand α within Field A = $\text{limit}_{t\to\infty} P(t)$.

$M(t,r,s)$ is the amount of Brand α purchased in period s beginning at time r after the first purchase of Brand α, aggregated over all buyers of Brand α, based on data available at time t.[1]

$E(t,r,s)$ is the amount of Product field A purchased in period s beginning at time r after the first purchase of Brand α, aggregated over all buyers of Brand α, based on data available at time t,

$R(t,r,s)$ is the repeat-purchasing rate for Brand α within Field A at time t, for purchasing interval s beginning at time r after the first purchase of Brand α.

$$= \frac{M(t,r,s)}{E(t,r,s)} \times 100,$$

(i.e., as a percentage)

$R(t,s)$ is the ultimate repeat-purchasing rate for Brand α within Product Field A at time t (based on a Purchasing interval s).
$= \text{limit}_{r \to \infty} R(t,r,s)$

$W(\alpha,t)$ is the amount of Field A purchased in period s beginning at time $t - s$, by all buyers of Brand α.

$W(A,t)$ is the amount of Field A purchased, in period s beginning at time $t - s$, by all buyers of Product field A.

$B(\alpha,A,t)$ is the buying rate factor for Brand α within Field A at time t.

$$= \frac{W(\alpha,t)}{W(A,t)},$$

(i.e., as a proportion).

Thus, given the three parameters P, $r(t,s)$, and $B(\alpha,A,t)$, the brand-share prediction model is the simple multiplicative model:

% brand-share prediction

$$= \frac{P\% \times R(t,s)\% \times B(\alpha,A,t)}{100}.$$

Figure 13A-1, a hypothetical example, illustrates the concept of brand-share prediction. In this example:

$$P = 34\%$$
$$R(s,t) = 25\%, \text{ and}$$
$$B(\alpha,A,t) = 1.20.$$

Although 34 percent of all buyers of Product field A have made at least one purchase of Brand α, they represent 41 percent of the total volume purchased of Product field A, buying at a rate 20 percent higher than the average Product field A buying rate. Thus, resulting only from the penetration of Brand α, the effective brand-share ceiling is 41 percent. This share may only be attained if all households making a first purchase of Brand α continue to buy only Brand α within Product field A. However, it is estimated that, having made a first purchase of Brand α, 25 percent of all subsequent purchases made in Product field A will be (ultimately) Brand α.

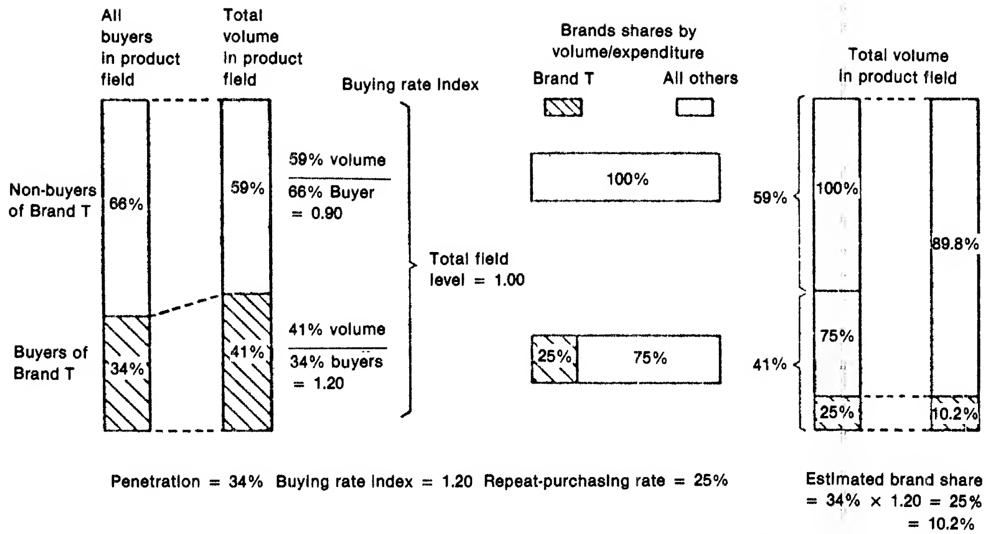


Figure 13A–1. Concept of Brand-Share Prediction

Therefore the estimated ultimate brand share is 25 percent of 41 percent, or 10.2 percent.

It is advantageous to produce repeat-purchasing rate analyses by segmented buying groups defined by the period of their first purchase of Brand α. The model is thus modified to become:

$$= \sum_{i=1}^{h} \frac{P_i\% \times R_i(t,s)\% \times B_i(\alpha A,t)}{100},$$

where $P_i\%$, $R_i(t,s)\%$, and $B_i(\alpha A,t)$ represent the ultimate penetration, repeat-purchasing rate and buying rate factor for the ith buying group, of which there are h.

Penetration Forecasts

When trying to formulate a penetration model, a family of growth models were considered and, based on several case histories, the modified exponential form proved the best choice,

Let

$$\Delta P(t) = \frac{P(t+1) - P(t-1)}{2},$$

K is the ultimate penetration,

a is the rate of growth parameter,

and

$\epsilon(t)$ is the random error associated with the measurements at time t.

Therefore, $\Delta P(t) = a(K - P(t)) + \epsilon(t)$.

This model is intuitively reasonable since it infers that the rate of increase of penetration at time t is proportional to the expected number of new buyers. The deterministic part of this stochastic model may be reduced to:

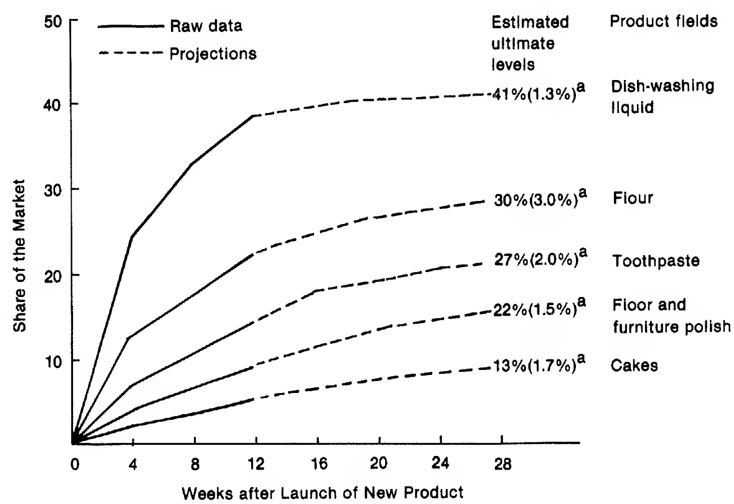
$$P(t) = K(1 - e^{-a \cdot t \cdot}),$$

where e is the exponential function.

It was suggested by Fourt and Woodlock [3] and Anscombe [1] described a method of obtaining the parameter's maximum likelihood estimates. However, the computations involved are arduous, and therefore a method has been devised applying the method of discounted least squares detailed in [4], to provide relatively quick estimates of a and K. As mentioned earlier the results of this work will be published soon.

Predictive Accuracy

However, it is obviously necessary to indicate the degree of accuracy with which the penetration model fits the data. In figure 13A-2 examples are



Source: Attwood Consumer Panel, Great Britain.

[a]100% = All households buying in specified field.

Figure 13A-2. Cumulative Penetration Model

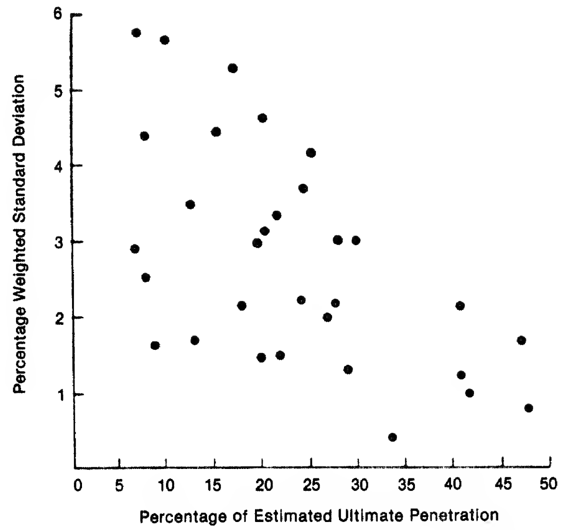


Figure 13A-3. Accuracy of the Fitted Model

given in which, based on the first 12 weeks' data, forecasts were produced for the next 16 weeks. Later, after receipt of subsequent observed data, it was possible to compare the predicted and actual data.

The critical measurement of predictive accuracy was defined as the percentage weighted standard deviation (p.w.s.d). For each example given, the 28-week p.w.s.d. is given in absolute product field. The p.w.s.d. at time t is defined as:

$$s(t) = \sqrt{\sum_{i=0}^{t} \frac{w^i}{W} \left(\frac{P(t-i) - \hat{P}(t-i)}{P(t-i)} \right)^2}$$

where

$P(t - i)$ is the $(t - i)$th (actual measurement of penetration),
$\hat{P}(t - i)$ is the $(t - i)$th (forecasted measurement of penetration),
w^i is the weighting function (w taken as 0.6),

and

$$W = \sum_{i=0}^{t} w^i.$$

Here $s(t)$ is a measure of the percentage deviation of the fitted model from the actual data giving more weight to the differences associated with the latest observations, i.e., those most critically affecting the estimate of K (see figure 13A-3).

Based on these results, it is possible to discriminate between relatively

stable penetrations and those affected by major marketing disturbances. By comparing the projected estimates of penetration, from the initial relatively stable state, with the actual penetration in the disturbance period, it is possible to quantify the effective increase or decrease in ultimate penetration and, with further analyses (repeat-purchasing rates, etc.) in eventual brand share.

Note

1. In all cases amount is weight, expenditure, volume, etc.; s is the selected length of the repeat-purchasing interval dependent on the frequency of purchase of buyers of Product field A.

References

1. F.J. Anscombe, "Estimating a Mixed-Exponential Response Law," *Journal of the American Statistical Association,* 56 (September 1961), 493–502.
2. J. Baum and K.E.R. Dennis, "The Estimation of the Expected Brand Share of a New Product," ESOMAR Congress, 1961.
3. Louis A. Fourt and Joseph W. Woodlock, "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* 25 (October 1960), 31–8.
4. W.G. Gilchrist, "Methods of Estimation Involving Discounting," *Journal of the Royal Statistical Society,* Series B, 29, No. 2, (1967), 355–69.

14 Prediction of New-Product Performance: An Analytical Approach

Henry J. Claycamp and
Lucien E. Liddy

Introduction

Individuals with the responsibility for planning the introduction of new consumer products are painfully aware of the complexities of the problems and the magnitude of the risks involved. Despite growing sophistication in the analysis of consumer needs and wants and vast investments in product research and development, most new products introduced by firms in the consumer package goods industry are commercial failures.

To improve this situation both business and academic researchers have recently devoted considerable effort developing techniques to help managers plan and control new product introductions. Although the primary objective of much of this activity has been the development of more reliable methods of predicting market acceptance levels, few attempts have apparently been made to build analytical models for predicting product performance before market introduction. With one notable exception [1] nearly all of the new product models discussed in the marketing literature are, in essence, techniques for forecasting equilibrium levels of consumer trials and repeat purchases from consumer panel data obtained during the initial phase of a test market or distribution rollout.[1]

Such techniques are valuable in identifying probable failures soon after launch. However, since marketing variables are not explicitly considered in making the forecast, they are of little value in making a priori evaluations of alternative introductory strategies or diagnosing the effects of individual elements of the marketing mix.

The purpose of this article is to present an analytical approach to the problem of predicting market responses to new consumer products which overcomes some of these limitations. In the following sections we will (a) describe the general structure of a model for predicting consumer trials and repeat purchases as a function of controllable and uncontrollable marketing

---

Reprinted from H.J. Claycamp and L.E. Liddy, "Prediction of New Product Performance: An Analytical Approach," *Journal of Marketing Research* 6 (November 1969): 414–420, published by the American Marketing Association.

variables and (b) present a detailed statistical analysis and empirical validation of the sector of the model used to predict initial trial levels.

The Model

The model presented here was designed to facilitate the planning and evaluation of alternative introductory campaigns for new consumer packaged goods. It was developed as part of a large-scale research project begun in 1965 by the advertising agency of N.W. Ayer & Son.

During the initial phase of this project, experienced marketing and advertising professionals were asked to state their operating assumptions about the way various controllable and uncontrollable variables influence consumer responses to new products. These statements and information gleaned from the marketing literature led to the specification of a conceptual model involving 14 "critical factors" and three types of consumer response to a given introductory campaign (see figure 14–1).

As seen from figure 14–1, the initial model was specified in terms of three interconnected submodels. The first submodel implies that the new product's position as expressed in the advertising with respect to existing products in the category (*PP*), the number (*AHI*) and quality (*CE*) of media advertising exposures and consumer promotions containing advertising messages (*CP**), and the level of consumer interest in the product category (*CI*) influence consumer behavior by generating knowledge about the new product and its advertised benefits. Since consumers may become aware of the new product as a result of exposures to it in retail outlets, "correct recall of advertising claims" (*AR*) rather than brand or advertising awareness is used as a measure of the direct impact of these variables.

The model's second sector implies that retail distribution factors (*DN*), packaging characteristics (*PK*), a known or family brand name (*FB*), the amount and type of consumer promotions (*CP*), the extent of consumer satisfaction with the product if it has been sampled (*PS**), the level of usage of the product category (*CU*), and the extent of consumer knowledge about the product, are directly associated with the level of initial purchases obtained by a new product during a given time period.

The final sector of the model implies that repeat-purchase levels will depend primarily on the extent of consumer knowledge, the level of initial trials, the price of the product relative to other products in the category (*RP*),[2] satisfaction with prior purchases (*PS*), and the frequency of purchase for products in the category (*PF*).

The three-part structure of the model makes it possible to obtain one direct and two indirect estimates of the impact of advertising variables on consumer response to a new product campaign, i.e., the effect of *PP, AHI,*

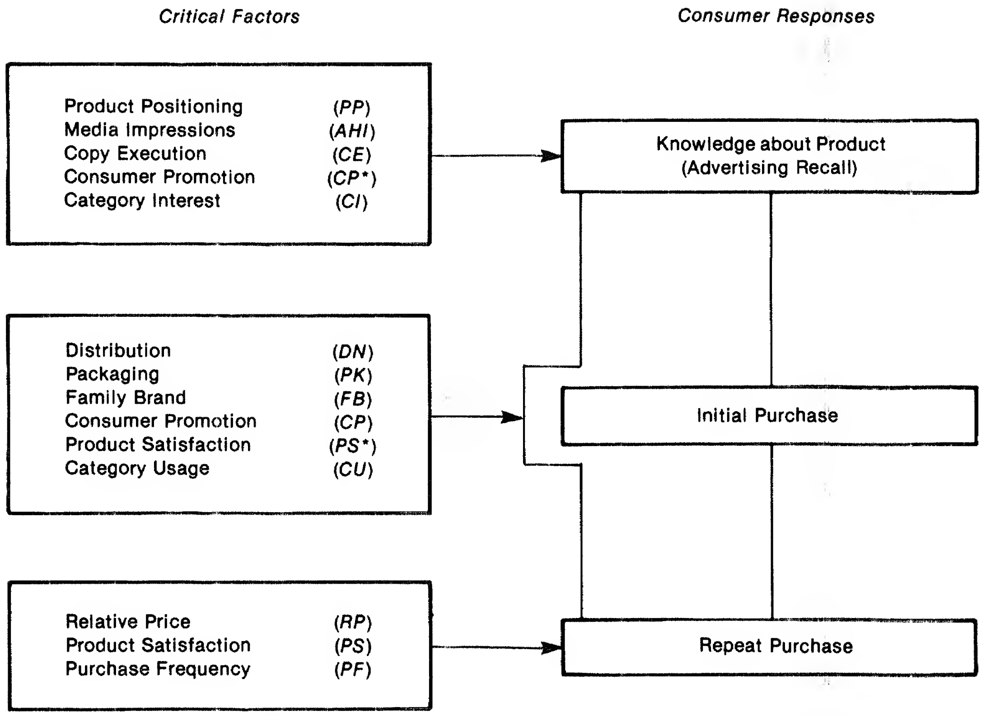


Figure 14–1. The Ayer New-Product Model

CE, and *CP** on advertising recall, and the effect of advertising recall on initial and repeat purchases. It also takes account of the fact that substantial numbers of initial and repeat purchases may be made by consumers who cannot recall the product's advertising messages.

Since the primary purpose of the introductory campaign is to inform consumers about the product and stimulate early trial, and most campaigns are planned in terms of 13-week cycles, a decision was made during the project's initial phase to concentrate on the first two sectors of the model shown in figure 14–1.

The Data

The data used in parameter estimation and model validation were collected between mid-1965 and early 1968. During this period approximately 60 new product introductions were monitored in the Philadelphia market area. Although the sample items were classified in 32 different product cate-

Table 14-1
Brand Names of Products Included in Estimation and Validation Sample

Brand Name	
Apple Jacks Cereal	Scope Mouthwash
Puppets Cereal	Reef Mouthwash
Quisp & Quake	Citrisun
Maxim Coffee	Adulton
Hills Bros. Coffee	Cope
Dole Pineapple-Pink Grapefruit Drink	Measurin
Del Monte Fruit Drinks	Vanquish
Knox Flavored Gelatin Drinks	New Kotex Plus
Moo Juice	True Cigarettes
Nestle's Quik Shake	Cigarette A
Great Shakes	Cigarette B
Start Instant Breakfast Drink	Cigarette C
Shake-a-Pudding	Petal Soap
Something Diff'rent	Phase III
Bounty Pudding	Hour After Hour Deodorant
Cool Whip	Bold Laundry Detergent
Chipnics	Cold Power
Sip-n-Chips	Crew
Jus'Fried Chicken	Cinch
Lipton Dinners	Palmolive Dishwashing Liquid
Honey Suckle	Pruf Spray Starch
Kraft Noodles Romanoff	Purex Super Bleach
Spaghetti-O's	Favor Furniture Polish
Pepperidge Farm Soups	Pronto Floor Care
Great American Soups	Epic Floor Wax
Kraft Frozen Sandwich Filling	Viva Towels
Hunt's Steakhouse Catsup	Handi-Wipes
Ultra Brite Toothpaste	Gaines Variety
Fact Toothpaste	Pet'm

gories, approximately 50 percent were food products. The remainder of the sample was made up of household supplies and personal care items. (A list of the brand names of the products used in the analysis is shown in table 14-1).

A detailed case history containing extensive information about characteristics of the introductory campaign, retail availability, and consumer responses was compiled for each product from a wide variety of data sources. For example, commercial data services and specially designed research projects were used to monitor television, radio, print and outdoor advertising, and to estimate the probable number of household exposure opportunities for each insertion.

Data on retail distribution, shelf space, in-store displays, and deals were obtained by auditing a sample of supermarkets and drugstores during the second and fourteenth week after the start of advertising for each new product.

Consumer surveys taken at the end of 13 weeks were used to estimate levels of advertising recall, initial and repeat purchases, and use of samples and coupons, product satisfaction, category usage, and category purchase frequency. Each consumer survey covered 250 housewives randomly selected from a total panel of 1,200 households. Although the total panel was established on a probability basis, it is not proportionately representative of households in the total Philadelphia market since over-sampling was done in younger, higher income areas.[3]

Quantitative values for variables such as product positioning, copy execution, and package quality were obtained by having a panel of experienced marketing and advertising executives rate each product, advertisement and package on predetermined scales. Subjective judgment was also used to develop weights for various kinds of consumer promotion, in-store displays, etc.

Starch "Ad Norm Scores" [6] were used as a measure of category interest for each product class.

The means and standard deviations of each of the variables used in the final model are shown in table 14–2.

Table 14–2
Means and Standard Deviations of Dependent and Independent Variables for Thirty-five New Products in Estimation Sample

Variable	*Description*	*Mean*	*Standard Deviation*
AR	Percentage of housewives able to accurately recall advertising claims at the end of 13 weeks	25.1	16.1
PP	Judged product positioning	35.5	15.3
AHI	Average number of media impressions/household	11.5	8.5
CE	Judged quality of advertising-copy execution	7.3	1.5
*CP**	Coverage of consumer promotion containing advertising messages adjusted for type of promotion	36.6	57.4
CI	Index of consumer interest in the product category	37.5	5.7
IP	Percentage of housewives making one or more purchases of the product during the first 13 weeks	13.9	11.9
$\hat{AR}$	Predicted advertising recall	25.1	13.8
DN	Retail distribution, adjusted for shelf space and special displays	58.5	14.8

Table 14–2 continued

Variable	*Description*	*Mean*	*Standard Deviation*
PK	Judged distinctiveness of package	0.556	0.132
FB	Known or family brand name	0.457	.051
CP	Coverage of consumer promotions adjusted for type and value of offer	63.9	55.1
*PS**	Index of consumer satisfaction with new-product samples	77.4	45.0
CU	Percentage of households using products in the category	67.0	23.2

Parameter Estimation

Methodology

The specific model for which parameters were estimated consists of two equations with two dependent and ten independent variables:

$$AR = a_1 + b_{11}(PP) + b_{12}(\sqrt{AHI^*CE}) + b_{13}(CP^*) + b_{14}(CI) + e_2. \tag{14.1}$$

$$IP = a_2 + b_{21}(\hat{AR}) + b_{22}(DN^*PK) + b_{23}(FB) + b_{24}(CP) + b_{25}(PS^*) + b_{26}(CU) + e_2. \tag{14.2}$$

Equation 14.1 is designed to predict the level of advertising recall to be expected 13 weeks after launch from the values of four independent variables. Since *CE* is a quantitative index of the quality of advertising exposures measured by *AHI,* both are combined in a single interaction variable. The square root transformation represents an assumption of diminishing returns to media weight.

Equation 14.2 is designed to predict the expected level of initial purchases at the end of 13 weeks as a function of the factors assumed to be directly or indirectly related to initial trials. For example, since the values of the independent variable $\hat{AR}$ are derived from 14.1, the parameter b_{21} can be interpreted as an estimate of the composite effect of the independent variables *PP, AHI, CE, CP*,* and *CI* on initial purchases, as well as an estimate of the nature of the relationship between advertising recall and initial trials. This formulation of the model makes it possible to use a pro-

cedure similar to two-stage least squares[4] to estimate the effects of all the variables on initial purchases while avoiding problems of multicollinearity and spurious correlation which occur if only one equation is specified.[5]

Stepwise least squares regression was used to estimate the parameters of the model from data on the introduction of 35 new products. Products in the estimation sample were randomly selected from a total of 58 for which data were available. The remaining 23 products were retained as a validation sample.

The parameters of equation 14.1 were estimated first and used to calculate values of $\hat{AR}$ for each of the 35 products. A second regression run was then made to obtain estimates of the coefficients in 14.2. The results of this analysis are shown in table 14–3.

Estimation Sample Results

Despite the use of subjective judgment to quantify several of the independent variables and the fact that many different kinds of products are included in the estimation sample, over 70 percent of the variance of both dependent variables is accounted for by the independent variables listed in table 14–3. The F ratios indicate that the correlations are highly significant

Table 14–3
Estimation-Sample Results

Variable	*Regression Coefficient*	*Standard Error*	R^2	*Sy*	$F(4.30)^a$
			Advertising Recall		
PP	.756	.122			
*AHI*CE*	2.122	.603			
*CP**	.039	.030			
CI	.392	.302	.725	8.970	19.790
(Intercept)	−35.876				
			Initial Purchase		
					$F(6,28)^a$
$\hat{AR}$	.370	.095			
*DN*PK*	.194	.107			
FB	9.245	2.414			
CP	.086	.024			
*PS**	.022	.033			
CU	.067	.059	.709	6.933	14.100
(Intercept)	−16.011				

[a] $p[F(4,30) \geq 4.02] = .01$; $p[F(6,28) \geq 3.53] = .01$.

despite the small sample size. Moreover, the signs of the regression coefficients are consistent with the hypotheses stated during the model's initial formulation, i.e., that each of the critical factors is positively related to the measures of consumer response.

The results of the regression analysis of 14.1 show highly significant relationships between levels of advertising recall for the 35 products in the estimation sample and the independent variables for product positioning and media advertising. Standardization of the regression coefficients indicates that these two variables are relatively far more important than CP^* and CI in the relationship, e.g., the Beta weights for PP, $\sqrt{AHI^*CE}$, CP^* and CI are .641, .411, .138, .138, respectively.

The regression statistics for $\hat{AR}$ and FB are particularly interesting. For example, the regression coefficient for $\hat{AR}$ is nearly four times as large as its standard error. Moreover, the magnitude of its β weight indicates that it is one of the most important variables in the relationship. (The β values for $\hat{AR}$, $DN{*}PK$, FB, CB, PS^*, and CU are .292, .179, .366, .374, .033, .194, respectively.) Since the values of $\hat{AR}$ are weighted sums of the independent variables in 14.1, this result supports the hypothesis that these variables exert an important influence on initial trial levels.

However, note that none of the independent variables in equation 14.1 has a statistically significant regression coefficient at the .05 level if they are used in place of $\hat{AR}$ in 14.2. This result, considered without the information shown in table 14–3, might lead one to conclude that initial trial levels are unrelated to product positioning and media advertising when in fact these variables are important determinants of $\hat{AR}$, and $\hat{AR}$ is one of the most important variables in the prediction of IP.

The regression coefficients for FB support the hypothesis that a known or "family" brand name on the new product has a positive net effect on trial purchases during the first 13 weeks of the campaign. The magnitude and significance levels of the coefficient are particularly surprising since FB is actually a dummy variable used as a proxy for the strength of the brand franchise (i.e, for products with a known brand name $FB = 1.0$, for all other products $FB = 0.0$).

It is also important to note that FB is virtually uncorrelated with the other independent variables in the relationship although one might expect family brand products to obtain higher levels of advertising recall and better retail distribution. (The zero-order correlation coefficients for FB with $\hat{AR}$ and DN are $-.13$ and .05, respectively).

Although the measure is crude, these results offer an important first approximation of the value of a "family brand umbrella" in the introduction of a new consumer product.

Since 14.1 and 14.2 were specified before data analysis, the results shown in table 14–3 can be interpreted as a meaningful test of the models'

validity. Although alternative functional forms and variable transformations were analyzed, none produced significant improvements over the initial model.

The real test of the model's validity, however, is its performance in predicting the dependent variables for a fresh set of products and its use as an operating tool. The results of these tests are discussed next.

Validation

Validation Sample Results

The regression coefficients shown in Column 2 of table 14-3 were used to predict the level of advertising recall and initial purchases for the 23 new products excluded from the regression analysis. Figures 14-2 and 14-3 show plots of the actual values of *AR* and *IP* obtained in consumer surveys and their corresponding predicted values.

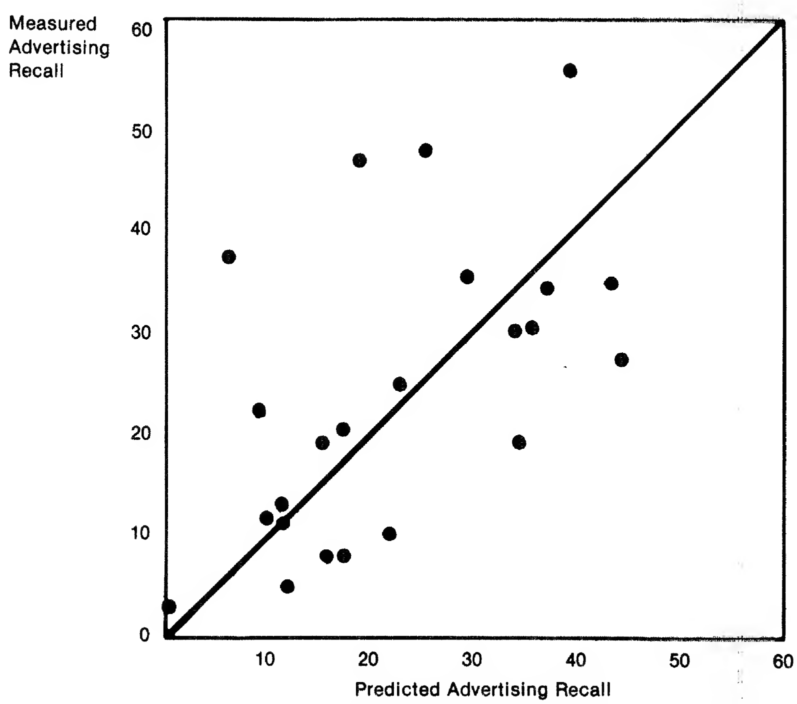


Figure 14-2. Measured and Predicted Advertising Recall Levels for Twenty-three New Products.

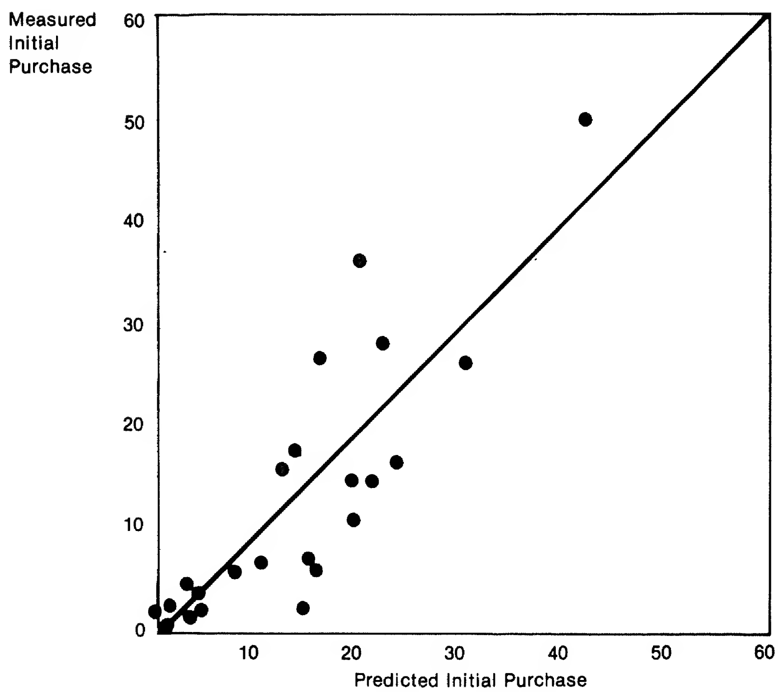


Figure 14–3. Measured and Predicted Initial Purchase Levels for Twenty-three New Products

It is clear from these diagrams that the model produces better predictions of initial purchases than advertising recall. For initial purchase, 13 of the predictions are within ±5 percentage points, and 20 are within ±10 percentage points of the actual level. For advertising recall, however, eight are within ±5 percentage points and 15 are within ±10 percentage points of the values measured in the consumer surveys. Product moment correlations between actual and predicted values provide a more striking comparison—the simple correlation coefficient for actual and predicted initial purchase is .95, and the corresponding statistic for actual and predicted advertising recall is .56.

The correlation statistic for advertising recall is, of course, greatly influenced by large errors made for four of the products. An analysis of the introductory campaigns for these products as well as products in the estimation sample for which similar errors were made indicates that advertising recall is typically underestimated when a "catchy" jingle is used in media advertising (e.g., one of the largest errors was found for "Spaghetti-O's").

Use of the Model

Since the model was designed to facilitate planning introductory campaigns, the most meaningful test is its usefulness in predicting, before launch, the probable outcome of a given campaign for a specific product.

The model has been used to predict initial purchase levels for seven different new products in eight widely dispersed geographic areas (figure 14–4). Although the parameters were estimated from data for new product introductions in Philadelphia, seven of the eight introductions occurred outside of the Philadelphia market area.

Most of the products were owned by companies other than Ayer clients. Some of the products were actually in the market place when the predictions were made. However, only prelaunch information was used to quantify the independent variables. Each company conducted its own consumer survey using standardized data collection instruments and sampling procedures to measure the actual trial levels achieved by its product by the end of 13 weeks.

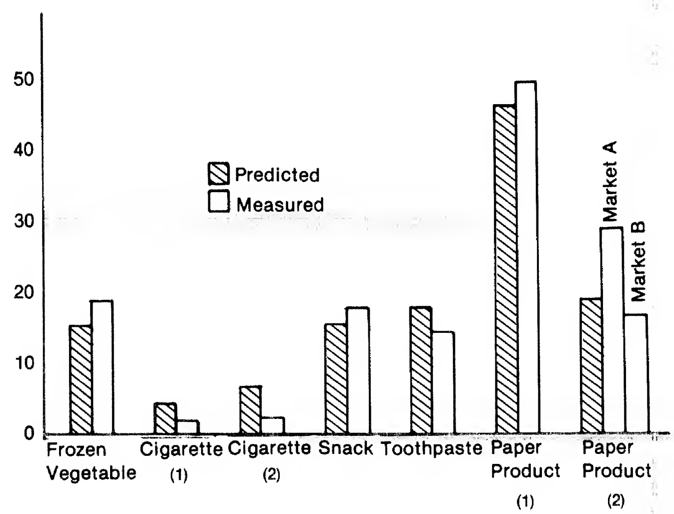


Figure 14–4. Measured versus Predicted Initial Purchase Levels for Seven New Products in Eight Market Areas

The results shown in figure 14–4 are remarkably consistent with those obtained in the analysis of the 23 products in the validation sample (figure 14–3) despite geographic dispersion of the markets and the fact that planned rather than measured values were used for the independent variables.

For example, in five of the eight cases the predicted value of initial purchases is within ± 5 percentage points of actual trial levels obtained in the consumer surveys, and only one prediction is off by as much as ten percentage points.

The largest discrepancy is for the paper product that was launched in two different market areas. Although the actual outcome is quite close in one of the markets, the measured trial level in the second market exceeds the predicted value by 13 percentage points. An ex-post analysis of this case revealed that the product achieved better than anticipated retail distribution in Market B.

Summary and Conclusions

The results of the regression analyses and the validation studies indicate that the model produces highly accurate predictions of the level of initial purchases to be expected 13 weeks after the start of a given campaign in a particular market area. Although the 13-week trial level may not be a perfect indicator of the ultimate success or failure of the product, it is widely used to measure initial market reactions and estimate long run penetration. The initial 13-week period also accounts for much of the risk and uncertainty associated with new product introductions.

The results also show that it is possible to produce reasonably accurate estimates of advertising recall levels from data which can be obtained before market introduction. They also provide considerable evidence that "correct recall of advertising claims" is a highly relevant measure of consumer response to advertising—at least for new products.

In fact, this measure provided the key to establishing a stable relationship between advertising variables and initial purchases. By regressing measured advertising recall on data describing the content, quality, and quantity of the advertising program and consumer interest in the product category, it was possible to obtain the values of a unique interaction variable—$\hat{AR}$—which had little correlation with the independent variables in 14.2 and a high degree of association with the level of initial purchases. Whereas, interaction variables are typically specified as the product of two or more independent variables, $\hat{AR}$ is the weighted sum of the variables in 14.1—regression analysis has been used to determine the weights.

Thus, the model evaluates the probable impact of a planned campaign on at least two important measures of new product performance.

Notes

1. For examples of these types of models, see [2, 4, 5].

2. Since most of the products of interest retail for less than one dollar, little risk is associated with a trial purchase. Hence, the primary impact of relative prices was assumed to be on repeat rather than purchasing behavior.

3. To check the nature of possible biases in the sample, consumer surveys were repeated for two products on restricted random samples of housewives not included in the original panel. No significant differences between panel and nonpanel statistics on advertising recall and initial purchases were found for either product.

4. Two-stage least squares would involve estimating the values of $\hat{AR}$ from the regression of AR on all the exogenous variables in the two equations. See [3, pp. 258–60].

5. A stepwise least squares regression analysis of IP on the independent variables in 14.1 and the last five variables in 14.2 was performed at one point in the analysis. Although the R^2 in this analysis was slightly higher than that shown in figure 14–3, multicollinearity among the independent variables caused the regression coefficients to fluctuate widely as variables were entered and deleted from the equation.

References

1. A. Charnes, *et al.,* "DEMON: Decision Mapping Via Optimum GO—NO Networks—A Model for Marketing New Products," *Management Science,* 12 (July 1966).
2. Louis A. Fourt and Joseph W. Woodlock, "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* 25 (October 1960).
3. J. Johnson, *Econometric Methods,* New York: McGraw-Hill Book Co., Inc., 1960, 258–60.
4. J.H. Parfitt and B.J.K. Collins, "The Use of Consumer Panels for Brand Share Predictions," *Journal of Marketing Research,* 5 (May 1968), 131–45.
5. William F. Massy, "Stochastic Models for Monitoring New Product Introductions," Chapter 4 in Frank M. Bass, Charles W. King, and Edgar A. Pessemier, eds., *Applications of the Sciences in Marketing Management,* New York: John Wiley & Sons, Inc., 1967.
6. Daniel Starch, Inc., *Starch 1968 Magazine Adnorms Report,* Mamaroneck, New York: Daniel Starch and Staff, 1968.

15 Structure of Repeat Buying for New Packaged Goods

Manohar U. Kalwani and
Alvin J. Silk

Sustaining a satisfactory level of repeat buying is widely recognized as the key to the successful launching of a new packaged good. The volume of ongoing sales ultimately realized for a new brand is determined by the number of "adopters" or repeat buyers and their purchase frequency. However, repeat purchasing develops gradually with attrition continuing to occur among customers even after they have bought a new brand several times. Tauber (1977, p. 80) suggests that "this 'wear-out phenomenon' is probably due to factors such as boredom after continual use, eventual resistance to price, or the consumers' need to use a product a number of times to be able to tell if it really fits their needs."

Knowledge of such adoption phenomena is the foundation for analytical efforts directed toward the assessment of new brands prior to test marketing (Silk and Urban 1978; Tauber 1977) and the subsequent forecasting of sales from early test market results. Building on the ideas put forth in Eskin's depth of repeat model (Eskin 1973), PANPRO, Eskin and Malec (1976) have recently reported progress in understanding the process of how repeat buying develops.

They propose a model of the evolution of the proportion of repeat buyers wherein the attrition of previous buyers occurs systematically from the first repurchase onward and is described by a simple decay function. Eskin and Malec (1976) further postulate that the frequency of purchasing remains unchanged after the first repeat. The key implication of these hypotheses is that the ongoing sales of a new packaged good can be obtained from a knowledge of the cumulative proportion of first repeaters and their purchase frequency (or equivalently, their average interpurchase time).

We report further analyses and empirical findings on the structure of repeat buying for new packaged goods. The theoretical rationale for a set of

---

Reprinted from M.U. Kalwani and A.J. Silk, "Structure of Repeat Buying for New Packaged Goods," *Journal of Marketing Research* 17 (August 1980):316–322, published by the American Marketing Association.

The authors are indebted to Jiri Nechleba and John McFadden for valuable assistance in data processing and analysis. Special thanks are due Gerald Eskin, Donald Morrison, and a *JMR* reviewer for helpful comments on the draft of the article. The assistance of the *Chicago Tribune* in supplying the consumer panel data analyzed is also gratefully acknowledged.

hypotheses suggested by the work of Eskin (1973) and Eskin and Malec (1976) is examined. The depth of repeat model is used to analyze the structure of repeat buying – i.e., the conversion of triers into the first repeat class, the conversion of first repeaters into the second repeat class, and so on for higher repeat levels. The objective is to compare the time paths of the cumulative proportions across various repeat levels. Do these penetration curves have similar functional forms? Are the interpurchase times approximately the same across repeat levels? Does the cumulative proportion of consumers who convert from one repeat class to the next increase with depth of repeat? Model parameters estimated by a maximum likelihood method are used to test hypotheses related to these questions.

In first section of the article we describe the depth of repeat model when time is treated in discrete intervals. We establish that the continuous analog of the discrete model is a familiar stochastic model based on a gamma mixture of exponential interpurchase times. A set of hypotheses suggested by Eskin (1973) and Eskin and Malec (1976) are presented next, and the likelihood expressions used to estimate model parameters are developed. Results form the empirical tests for four new products – two toothpastes and two coffees—are examined, and the implications of these results for consumer research on new packaged goods are discussed.

Depth of Repeat Model

Our analysis of repeat buying follows the work of Fourt and Woodlock (1960), Massy (1969), and Eskin (1973). Depth of repeat classes are defined as the penetration or cumulative proportion of consumers who repeat a Jth time ($J = 1,2,3\ldots$) given that they had previously made $J - 1$ repeat purchases. Note that the repeat class (or level) $J = 1$ refers to first repeaters, the repeat class $J = 2$ refers to second repeaters, and so on for higher repeat levels.

Two postulates underlie the penetration model proposed by Fourt and Woodlock (1960). First, there is a ceiling on the proportion of consumers who convert from one repeat class to the next. Second, the number of consumers who enter the next repeat level in each time period is a constant fraction of the remaining consumers who will eventually convert into the next class. Fourt and Woodlock (1960, p. 32) cite empirical evidence to support both of these postulates.

> Observation of numerous annual cumulative penetration curves shows that (1) successive increments to these curves decline, and that (2) the cumulative curves seem to approach a limiting penetration of less than 100 per cent of households—frequently far less.

Consider repeat level J where $J = 1,2,3, \ldots$. Given the preceding two postulates it follows that the proportion of consumers, $P(L)$, who convert into the modeled repeat level J during the Lth time period is given by:[1]

$$P(L) = F(L) - F(L-1) = (1-\rho)(\alpha - F(L-1)) \quad (15.1)$$

$L = 1, 2, 3, \ldots$

where $F(L)$ is the cumulative proportion of consumers who convert into a particular repeat level, J, by the Lth time period since their previous purchase; ρ is a constant; and α denotes the cumulative proportion who will eventually convert into the repeat class J. Thus, α represents the ceiling on penetration. Solving equation 15.1, recursively (with $F(0) = (0)$), one obtains:

$$P(L) = (1-\rho)(\alpha - F(L-1)) = \alpha\rho^{L-1}(1-\rho) \quad (15.2)$$

$L = 1, 2, 3, \ldots.$

$$F(L) = \sum_{L=1}^{L} \alpha\rho^{L-1}(1-\rho) = \alpha(1-\rho^L) \quad (15.3)$$

$L = 1, 2, 3, \ldots.$

Figure 15-1 is a graphic display of the penetration function, $F(L)$, over time. An examination of the preceding expressions for $P(L)$ and $F(L)$ reveals that the purchasing behavior of consumers who enter repeat class J is given by a geometric distribution. The average interpurchase time, τ, of these consumer who will eventually convert into some repeat level J is given by:[2]

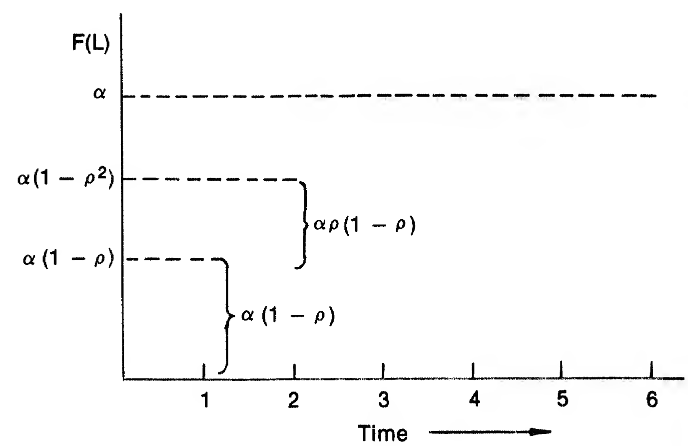


Figure 15-1. Penetration for Repeat Level *J*

$$E[L] = \tau = \frac{\sum_{L=1}^{\infty} L\alpha\rho^{L-1}(1-\rho)}{\sum_{L=1}^{\infty} \alpha\rho^{L-1}(1-\rho)} = \left(\frac{1}{1-\rho}\right) \tag{15.4}$$

which can be rearranged as:[3]

$$\rho = \left(\frac{\tau - 1}{\tau}\right) \tag{15.5}$$

Note that the expression for interpurchase time given in equation 15.4 is based on the assumption that data on conversions into the Jth repeat level are available over an infinite period of time. In practice, however, the period over which conversion rates are observed is bound to be limited. If we assume that the conversion into the Jth repeat class is observed for T periods, the expression for average interpurchase time, τ_T, becomes:

$$\tau_T = \frac{\sum_{L=1}^{T} L\,\alpha\rho^{L-1}(1-\rho)}{\sum_{L=1}^{T} \alpha\rho^{L-1}(1-\rho)} = \left(\frac{1}{1-\rho}\right) - \left(\frac{T\rho^T}{1-\rho^T}\right) \tag{15.6}$$

Hence, the average interpurchase time, τ_T, for truncated data is smaller than the theoretical average interpurchase time, τ.

To this point the consumers entering a particular repeat level have been assumed to be homogeneous. Empirical evidence, however, indicates that the early entrants into a repeat class tend to be heavier buyers of the product than later entrants (Eskin 1973; Fourt and Woodlock 1960; Massy 1969; Parfitt and Collins 1968). Fourt and Woodlock (1960, p. 33–4) note that the estimated penetration levels based on equation 15.3 fit the data well except that the predictions for distant time periods tend to be too low. The poor fit is attributed to heavier buyers converting into a repeat class earlier than light buyers. The addition of a trend factor, δ, provides for this "stretch out" of penetration, which leads to the following adjustments in equations 15.2 and 15.3:

$$P(L) = \alpha\rho^{L-1}(1-\rho) + \delta \quad L = 1, 2, 3, \ldots. \tag{15.7}$$

$$F(L) = \alpha(1-\rho^L) + \delta L \quad L = 1, 2, 3, \ldots. \tag{15.8}$$

This modification allows the ceiling to be a linear function of time instead of a fixed quantity. The expression for average interpurchase time given a finite observation time period, T, now becomes:

$$\tau_T^\delta = \frac{\sum_{L=1}^{T} L(\alpha\rho^{L-1}(1-\rho)+\delta)}{\sum_{L=1}^{T}(\alpha\rho^{L-1}(1-\delta)+\delta)}$$

$$= \frac{\alpha\left\{\dfrac{1-\rho^T}{1-\rho} - T\rho^T\right\} + \dfrac{\delta T(T+1)}{2}}{\alpha(1-\rho^T)+\delta T}$$

Empirical evidence presented by Fourt and Woodlock (1960) and Eskin (1973) suggests that the numerical value of the "stretch out" factor, δ, is very small in comparison with the conversion proportion term, α. Therefore, the effect of the δ term in equation 15.9 is likely to be minor, especially when the observation period over which the purchase data are available is of limited duration (say, 12 or 24 weeks).

The continuous analog of the discrete depth of repeat model is the familiar NBD model (Ehrenberg 1959) which assumes exponentially distributed interpurchase times (or equivalently, a Poisson distribution of purchase events across successive time periods of equal length) for an individual household, and gamma heterogeneity over the population.

Following the NBD model, the density function of time to conversion into the Jth repeat class is distributed as a negative exponential; hence:

$$f(t) = \lambda e^{-\lambda t} \qquad \lambda > 0 \tag{15.10}$$

If we assume that the mixing distribution of λ is gamma,[4] with μ as the scale parameter and v as the shape parameter, the expressions for $f(t)$ and $F(t)$ are

$$f(t) = \frac{v}{\mu}\left(\frac{\mu}{\mu+t}\right)^{v+1} \tag{15.11}$$

and

$$F(t) = 1 - \left(\frac{\mu}{\mu+t}\right)^{v} \tag{15.12}$$

A refinement of this model allows for a "zero group" of consumers who are not in the market for the new product and are to be excluded from the "relevant population." Kalwani and Silk (1978) used a likelihood ratio test to determine the desirability of this refinement. The unconstrained model which allows for the "zero group" provided a fit superior to that of the

constrained model in each of the four cases tested. The expressions for $f(t)$ and $F(t)$ with this modification are given by:

$$f(t) = A\,\frac{v}{\mu}\left(\frac{\mu}{\mu + t}\right)^{v+1} \tag{15.13}$$

and

$$F(t) = A\left\{1 - \left(\frac{\mu}{\mu + t}\right)^{v}\right\} \tag{15.14}$$

Penetration estimates from equation 15.14 are compared hereafter with those from the discrete case (equation 15.8) to show whether the latter model (which is simpler and easier to interpret) performs as well as the continuous model.

Hypotheses on the Structure of Repeat Buying

For the depth of repeat model set forth, the cumulative proportion of consumers who convert into the Jth repeat level within L periods of previous purchase is given by

$$F_J(L) = \alpha_J(1 - \rho_J^L) + \delta_J L \qquad L = 1, 2, \ldots \tag{15.15}$$

where the suffix J denotes the repeat class and the time variable L is measured in time periods since the last purchase. As indicated before, one purpose here is to compare the penetration curves represented by equation 15.15 across various repeat levels. How do the parameters α, ρ, and δ vary with J? Eskin (1973) proposed three hypotheses about the patterns in α, ρ, and δ across repeat levels and tested them with purchase data for six *established* products.

Eskin's first hypothesis was that for a new product the parameter ρ_J has approximately the same value across all the repeat levels. This implies that the average interpurchase times are approximately equal across all repeat classes. The average interpurchase time for a particular new product is bound to be larger than the average interpurchase time for the product class which, of course, depends on the size of the offering. Also, the average interpurchase time depends on the proportion of consumers who are completely loyal to the best brand. The larger the proportion of buyers who are committed to the new product and buy it every time, the smaller will be the average interpurchase time. In testing this hypothesis with data for six

established products, Eskin (1973, p. 127) found that ρ_J "does not fluctuate excessively nor does it exhibit a strong trend."

A second hypothesis put forth by Eskin (1973) states that δ_J takes on the same value across all repeat levels. In this case, Eskin found that δ_J's "vary in a relatively small range but do tend to exhibit a negative trend over J." Both Eskin (1973) and Fourt and Woodlock (1960) found that the magnitude of the δ term is small across all repeat levels. In other words, in the expression for penetration at the Jth repeat class given by equation 15.15, the first term—$\alpha_J(1 - \rho_J^L)$—forms the major contribution, especially at higher repeat levels.

A third hypothesis suggested by Eskin (1973) applies to the conversion proportion terms, α_J's. Eskin postulated that α_J's could be obtained from a geometric distribution of the form:

$$\alpha_J = \alpha_\infty(1 - \gamma^J) \qquad J = 2, 3, \ldots \tag{15.16}$$

where the limit, α_∞, is usually unity or slightly less than unity. He found that the relationship given by equation 15.16 provided a good fit to the data, with the estimates of α_∞ being less than unity for each of the six established products studied. When equation 15.16 was fitted to a data base that consisted of observations for an unspecified cross-section of new products, Eskin and Malec (1976) obtained an estimate of 0.626 for the parameter γ. With $\alpha_\infty = 1$, this value implies that $\alpha_2 = 60\%$, $\alpha_3 = 74\%$, $\alpha_4 = 84\%$, $\alpha_5 = 90\%$, etc.

The systematic pattern in conversion proportion terms across a variety of new packaged goods can be interpreted as follows. Consumers will try a new product on the basis of its expected performance. The first or trial purchase of a new product is ordinarily followed by numerous use experiences in which a consumer evaluates the product quality. The new product is repurchased (generally to the exclusion—partial or complete—of other products that the consumer was previously buying) only if the consumer is satisfied with the test product in comparison with previously used products. Commitment to the new product increases as it is repurchased again and again, and ultimately the consumer "adopts" the product.

What are the practical implications of these three hypotheses? The first hypothesis implies that, for a given new product, the mean interpurchase times (or equivalently, the parameter ρ_J) of second, third, fourth, repeaters are the same as those of first repeaters. The second hypothesis states that for a new product the parameter δ_J is constant across repeat levels. The final hypothesis on the systematic pattern in the values of $\alpha_2, \alpha_3 \ldots$ is applicable *across* new packaged goods.

If we assume that the magnitude of δ_J is small across repeat levels, the

hypotheses taken together imply that for a new product the penetration for various classes of repeat (see equation 15.1) can be obtained from knowledge of the average interpurchase times of first repeaters (or equivalently, ρ_1) and the cumulative proportion of first repeaters, α_1. Further, to the extent that the interpurchase time of a new brand resembles those of existing brands in the product class, the cumulative repeat proportions for various repeat classes can be estimated solely from the cumulative proportion of first repeaters.

Parameter Estimation

Maximum likelihood methods were used to estimate model parameters for each repeat level. It is well known (see Rao 1965) that under very general regularity conditions, maximum likelihood estimates are best asymtotically normal (BAN). That is, they are consistent, asymptotically normal, and asymptotically efficient. In addition, maximum likelihood estimates are invariant. Eskin (1973) employed a least squares approach to estimate model parameters. The fact that the dependent variable in the regression equation is defined as a cumulative proportion, and hence is nondecreasing, makes the presumption of uncorrelated error terms tenuous. Detecting autocorrelation, Eskin (1973) applied a generalized least squares procedure to obtain efficient parameter estimates. In some preliminary analyses of the products studied here, the same generalized least squares approach followed by Eskin was investigated but the results were unsatisfactory. Estimates of α that exceeded unity were found for some repeat levels, and some estimates of δ were negative. Measurement errors in the penetration data might account for these difficulties. The estimates were also unstable because of small sample sizes for two of the new products considered, especially at higher repeat levels.

The difficulty with using the maximum likelihood method is that it is not possible to obtain closed-form analytical solutions. Therefore, numerical optimization is required to obtain maximum likelihood estimates. For this purpose, we used a general optimization procedure developed by Kalwani (1975). The computer program (written in FORTRAN IV) executes an accelerated pattern search technique. The search can be limited to the feasible range of solutions, but it was not necessary to impose such restrictions for any of the four new products investigated.

The likelihood expression that is maximized is determined by the following procedure. Consider any one of the repeat levels, J. Let n_L denote the number of consumers who enter repeat level J during the Lth time period since their previous purchase. Furthermore, let $\bar{n}_L$ denote the number

of consumers who had L time periods available to convert into the Jth repeat level but did not do so. Then $\Sigma_{L=1}^{52} n_L = n$ represents the total number of buyers who convert into repeat level J within 52 weeks of their entry into the $(J - 1)$th repeat class. Similarly, $\Sigma_{L=1}^{52} \tilde{n}_L = (m - n)$ denotes the total number of consumers who have not entered the repeat class J given that, altogether, m consumers made $(J - 1)$ purchases of the new product.

The likelihood expression given purchase data n_L and $\tilde{n}_L$ for, say, $L = 1, \ldots 52$, is given by

$$l(n_L' \, s, \tilde{n}_L' \, s; \alpha, \rho, \delta)$$

$$= \prod_{L=1}^{52} (P[L])^{n_L} \prod_{L=1}^{52} (1 - F[L])^{\tilde{n}_L} \tag{15.17}$$

where

$$P(L) = \alpha\rho^L(1 - \rho) + \delta,$$

$$F(L) = \alpha(1 - \rho^L) + \delta L.$$

The left multiplicand in equation 15.17 represents the joint probability of finding n_L $(L = 1, 2, \ldots, 52)$ consumers who converted into the Jth repeat class within L time periods of their previous purchase. Similarly, the right multiplicand forms the joint probability of observing that $\tilde{n}_L(L = 1, 2, \ldots, 52)$ consumers failed to enter repeat class J within L time periods of their previous purchase.

Findings

Data Base

The results reported are based on an analysis of purchase data for four new products, two brands of toothpaste—Ultra Brite and Plus White—and two brands of freeze-dry coffee—Maxim and Taster's Choice. The source of purchase information is the panel data collected by *Chicago Tribune's* Family Survey Bureau, which contains purchase records for 530 households. The largest amount of purchase information was available for Ultra Brite; the observations covered a three-year period after its introduction. For Plus White and the two brands of coffee, the purchase data extended over a period of about two years after introduction of each of the products. Findings are reported only for those repeat levels where the sample sizes are at least 30.

Results

Maximum likelihood estimates of parameters, displayed in table 15–1 to 15–4, are the basis for evaluation of the three aforementioned hypotheses.

The first hypothesis is that the parameter, ρ, would be approximately, the same across repeat levels for each of the four products. For Ultra Brite, which has the largest sample size, table 15–1 reveals very little variation across the four repeat levels, with the estimated ρ's fluctuating around .94. The results for Plus White (table 15–2) and the two coffees (tables 15–3 and 15–4) appears less consistent, but the variability is still slight. The ranges of the estimated ρ's for these three products are: Plus White, .89–.93; Maxim, .84–.92; and Taster's Choice, .88–.92. Note, however, that some variation in ρ is expected for the two coffees because ρ is related to the interpurchase time, which is affected by variations in the consumption of coffee.

The second hypothesis is that the parameter δ would be approximately the same across repeat levels for each of the four products. Results show that the estimated values of δ are small, generally less than 0.2%, which is consistent with the empirical experience of Fourt and Woodlock (1960).

The final hypothesis concerns the conversion proportion terms, α_j's. On the basis of the empirical results reported by Eskin and Malec for a cross-section of new products, we postulated that these terms would increase systematically to α_∞, which is expected to be slightly less than unity. More specifically, given Eskin and Malec's estimate of $\gamma = .636$ for equation 15.16, we expected that the α_j's would take the following values: $\alpha_2 = 60\%$, $\alpha_3 = 74\%$, $\alpha_4 = 84\%$, etc.

The maximum likelihood estimates of the α_j terms for the four new products are shown in table 15–1 through 15–4. Though the specific estimates of the α_2, α_3, and α_4 for Ultra Brite in table 15–1 of the α_j's are not exactly equal to the hypothesized values, they do exhibit a nondecreasing pattern. Table 15–5 shows the differences between estimated and hypothesized values of the α_j's. Note that the deviations are all within approximately 10% of the hypothesized values.

Finally, we can compare penetration estimates from the unconstrained version of the continuous model with those from the discrete model. Ultra Brite is chosen to illustrate the results because it is the brand for which the largest amount of purchase information was available for estimation.

Table 15–6 displays the penetration for time intervals (since last purchase) of 12, 24, 36, and 52 weeks. The estimated penetrations based on the discrete and unconstrained versions of the continuous models are very close to the actual penetration levels, particularly for the 52-week interval. (Note that the results shown in table 15–6 relate to the goodness of fit rather than the predictive accuracy of the discrete and continuous models.) These results indicate that the discrete model (which has parameters that are easier

Table 15–1
Ultra Brite Toothpaste

Repeat Level	*Sample Size*	*Parameter estimates* α	ρ	δ
1	134	.5410	.9394	.0001
2	86	.6786	.9474	.0000
3	66	.7116	.9446	.0018
4	49	.7560	.9363	.0000

Table 15–2
Plus White Toothpaste

Repeat Level	*Sample Size*	*Parameter Estimates* α	ρ	δ
1	64	.3929	.9264	.0000
2	39	.5378	.8928	.0014
3	31	.7991	.9042	.0000

Table 15–3
Maxim Coffee

Repeat Level	*Sample Size*	*Parameter Estimates* α	ρ	δ
1	118	.3027	.8394	.0025
2	75	.6191	.8837	.0014
3	54	.7605	.9190	.0000

Table 15–4
Taster's Choice Coffee

Repeat Level	*Sample Size*	*Parameter Estimates* α	ρ	δ
1	150	.6425	.9235	.0000
2	99	.6694	.8852	.0014
3	72	.8092	.8755	.0003

to interpret) provides a fit as good as that of the unconstrained continuous model. The two key parameters of the discrete model are ρ and α. The parameter ρ is directly related to average interpurchase time, and the parameter α simply represents the cumulative proportion of consumers who convert from a given repeat level to the next.

Discussion and Conclusions

Results are reported from some empirical tests of three hypotheses set forth by Eskin (1973) and Eskin and Malec (1976) on the structure of repeat buying. Maximum likelihood estimates of model parameters were developed for four new packaged goods—two toothpastes and two coffees—and compared for consistency with the hypothesized patterns. The first hypothesis predicts that for a given new product the parameter ρ_J, a measure of average interpurchase time, would be constant across repeat levels. Overall, the parameter estimates obtained are found to be consistent with the hypothesis—especially in the case of Ultra Brite and to a lesser but still supportive degree for the other three new brands. As hypothesized, the parameter δ_J displays little variation across repeat levels and further, its magnitude is generally found to be very small.

The empirical tests of the final hypothesis on systematic patterns in cumulative conversion proportion terms α_2, α_3 . . . reveal that, although estimates of these parameters deviate somewhat from their hypothesized values, they do exhibit the postulated nondecreasing pattern (see table 15-5 for differences between estimates and hypothesized values of α_J's).

Though these results are generally encouraging, attention needs to be drawn to the limited number of repeat levels used to test the hypotheses. In spite of the fact that the purchase records for each of the four products

Table 15-5
Differences Between Estimates and Hypothesized Values of α_J's

		Observed Value − Hypothesized Value			
Parameter	*Hypothesized Value*	*Ultra Brite*	*Plus White*	*Maxim*	*Taster's Choice*
α_2	.60	.0786	−.0622	.0191	.0694
α_3	.74	−.0284	−.0409	.0205	.0692
α_4	.84	−.0840	n.a.	n.a.	n.a.

n.a. denotes "not available."

Table 15-6
Observed and Estimated Penetrations for Ultra Brite
(percent)

Repeat Level	T = 12 Weeks			T = 24 Weeks			T = 36 Weeks			T = 52 Weeks		
		Estimated			Estimated			Estimated			Estimated	
	Observed	Discrete	Continuous	Observed	Discrete	Continuous	Observed	Discrete	Continuous	Observed	Discrete	Continuous
1	30.3	28.7[a]	28.8	43.3	42.2	42.2	48.8	48.8	48.8	52.8	52.5	52.8
2	32.8	32.4	32.6	50.7	49.3	49.2	58.2	58.2	58.1	64.2	63.8	64.2
3	36.0	37.4	37.4	59.3	57.4	57.2	67.4	68.5	68.4	76.7	76.8	76.7
4	43.3	41.3	41.5	62.7	60.0	59.8	68.7	68.5	68.2	73.1	73.1	73.1

[a] A total of 28.7 percent of first purchasers entered the first repeat class within twelve weeks of trying the new product.

extended over a period of at least two years, the small size of the *Chicago Tribune* panel yielded only a few repeat levels with sample size of 30 or more.

The aforementioned hypotheses carry important implications for the tasks of making either pretest market or early test market forecasts of the time path and equilibrium level of penetration for new brands. Given that these hypotheses about repeat buying hold, it follows that penetration levels for various repeat classes depend primarily on two factors: the average interpurchase time of first repeaters (or equivalently the paremeter ρ_1) and the cumulative proportion of first repeaters. A reasonable initial estimate of ρ_1 (or τ_1) could be obtained for a new brand by examining the interpurchase times of existing brands in the product class (taking into account any differences in package sizes). Then the repeat sales for the new brand could be forecast from a knowledge of the cumulative proportion of triers who repeat buy it at least once. Research is underway to link these hypotheses about buying to measurement methods that are used to assess new brands prior to test marketing (Silk and Urban 1978).

Notes

1. Note that in equation 15.1 the number of consumers purchasing the new packaged good (or the number of buyers entering a particular repeat class) is not influenced by the number who have already purchased the new product (i.e., $F(L - 1)$).

2. The expression for variance can be obtained in a manner similar to the derivation of expected value of L.

$$\mathrm{VAR}[L] = \frac{\sum_{L=1}^{\infty} L^2 \alpha \rho^{L-1}(1 - \rho)}{\sum_{L=1}^{\infty} \alpha \rho^{L-1}(1 - \rho)} - (E[L])^2 = \frac{\rho}{(1 - \rho)^2}$$

3. Gerald Eskin (personal correspondence) has informed us that equation 15.5 corrects a typographical error that appears in equation 4 of Eskin and Malec (1976, p. 231).

4. This is the key difference between the discrete and continuous models. In the discrete case, the "stretch out" factor is used to allow for the heterogeneity between early and late entrants into a repeat class. The use of the gamma distribution in the continuous case allows more flexibility in describing consumer heterogeneity.

References

Ehrenberg, A.S.C. (1959), "The Pattern of Consumer Purchases," *Applied Statistics,* 8, 26–41.

Eskin, Gerald J. (1973), "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model," *Journal of Marketing Research,* 10 (May), 115–29.

Eskin, G.J., and John Malec (1976), "A Model for Estimating Sales Potential prior to the Test Market," *Proceedings,* Educators Conference, Series No. 39. Chicago: American Marketing Association, 220–33.

Fourt, Louis A., and Joseph W. Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* (October), 31–8.

Kalwani, M.U. (1975), "Zero-Order Models: Specification, Data Requirements, and Estimation," unpublished Ph.D. dissertation, Graduate School of Business, Columbia University.

Kalwani, M.U., and Alvin J. Silk (1978), "Structure of Repeat Buying for New Packaged Goods," Working Paper No. 1032–78, Sloan School of Management, M.I.T. (October).

Massy, William F. (1969), "Forecasting the Demand for New Convenience Products," *Journal of Marketing Research,* 6 (November), 405–12.

Parfitt, J.H., and B.J.K. Collins (1968), "Use of Consumer Panels for Brand-Share Prediction," *Journal of Marketing Research,* 5 (May), 131–45.

Rao, C.R. (1965), *Linear Statistical Inference and Its Applications,* New York: John Wiley & Sons, Inc.

Silk, Alvin J., and Glen L. Urban (1978), "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research,* 15 (May), 171–91.

Tauber, Edward M. (1977), "Forecasting Sales Prior to Test Market," *Journal of Marketing,* 41 (January), 80–4.

16 TRACKER: An Early Test-Market Forecasting and Diagnostic Model for New-Product Planning

Robert Blattberg and
John Golanty

Introduction

A new product forecasting model is described that predicts year-end test market sales from early test market results (usually three months). In addition to offering a sales forecast, the model is designed to provide diagnostic information about the product's performance. By relating actual sales to advertising expenditures, price, product quality, and consumer data, the model indicates how a marginal or unsuccessful product can be redesigned or the marketing mix changed to increase its success.

The model also can be used for new product planning. Given a media plan, price, sampling level, couponing, and some estimate of repeat use, a pre-test market forecast of year-end sales can be made which allows management to alternate different marketing plans to see which best meets profit or sales goals. Obviously, the pre-test market forecasts are based solely on "norms" and not actual test market results. When the product is introduced results may fall below (or above) norms, but before test market the model will indicate "average" responses to different media and pricing strategies.

Many other new product forecasting models are available [see 2, 4, 7–14]. The model presented here has several distinct features. First, it requires only three months of test market data whereas many other models require longer time periods (commonly six months or longer). Second, it uses survey data. Almost all other models require panel data which are much more expensive and take longer to recover and process from the marketplace. Third, the model is inexpensive to use. The total cost of using

---

Reprinted from R. Blattberg and J. Golanty, "TRACKER: An Early Test-Market Forecasting and Diagnostic Model for New-Product Planning," *Journal of Marketing Research* 15 (May 1978):192–202, published by the American Marketing Association.

The authors thank James Livingston, David Olson, and Subrata Sen for their helpful comments. This research was funded in part by National Science Foundation grant GS-40033.

the model, including data collection, is roughly $15,000. Many other new product models offering comparable forecasts cost $50,000 to $100,000. Fourth, the model is very easy to understand and to use. (To aid users, a time sharing version is available.) Many of the other new product forecasting models are not applied because very few people understand them. Since the primary purpose was to develop a model that would be used, the authors believed that the model should be kept as simple as possible. Therefore, the model uses only one or two variables at each stage, all easily understandable, and the relationships involved can be explained to any potential user. Fifth, the model is very accurate. TRACKER has been applied to 11 new product introductions since it was developed.[1] The forecasts have been made with three months of test marketing in most cases. Table 16–1 lists the results. It should be noted that the forecasts given in table 16–1 are not merely fitted values from a model, but are actual forecasts. Finally, the model offers its users diagnostic information about problems with respect to awareness, trial and repeat rates, and can indicate causes of these problems, e.g., low spending, high price, poor product quality. All of this diagnostic information is based on the standard data collected. Other models, particularly models using panel data, often require additional studies to identify the causes of a new product's failure. The model forecasts and simultaneously diagnoses product problems.

Overview of the Model

The basic theory underlying the model is similar to that outlined in [6]. The process begins with advertising of the new product. Advertising results in a fraction of potential users becoming aware of the product and the rest entering the nonaware class. The aware consumers then become triers or they enter the nontrier class. The members of the trier class develop attitudes about product quality and product satisfaction from their initial trial experience. These attitudes determine whether they become repeat users or move into the nonrepeat user class. Finally, repeat users can either continue to repeat purchase or move into the nonrepeat users class. Figure 16–1 is a flowchart of the model.

The data inputs for the model are obtained from questionnaires administered to potential users of the product. Table 16–2 lists the information available from the questionnaire. Three waves of 500 to 1,000 questionnaires are collected for each new product introduction. The waves are sequenced every four weeks after introduction.[2] Also collected for each product are mediaweight, price, special distribution problems, and other data which may be applicable to a specific introduction.

In estimating parameters for each model, separate estimates are com-

Table 16-1
Validation of Projections

Case	*Forecast Made After*	*Forecast*	*Actual Results*
1	4 months	1.6 percent share	2.0 percent share
2	2 months	0.9 percent share	1.0 percent share
3	2 months	0.3 percent share	Withdrawn after 3 months with 0.3 share
4	7 months	$36 million annual rate	$39 million annual rate at 15 months
5	3 months	Early product failure	Withdrawn—6 months
6	3 months	2.1 percent share	2.3 percent share
7	2 months	Early product failure	Withdrawn—3 months
8	3 months	5.8 percent share	5.2 percent share after 8 months
9	7 months	$25 million rate	$23 million annual rate at 15 months
10	4 months	$7.5 million first year	Withdrawn at 5 months
11	6 months	30 percent share	26 percent share

puted when possible for each product category (e.g., dog food, cereal, etc.)[3] This step is very important for making predictions and using the model for planning. There is large variation in parameter estimates across product categories. For example, some product categories (not brands) are very responsive to advertising, others are not; some have high trial among aware consumers, others do not. Therefore, separate parameter estimates are needed for each product category. An example is given in the following section.

The Awareness Model

The Model

In developing the advertising-awareness model, several prior ideas about how advertising affects the awareness were considered. The goal was to develop a model consistent with both the limitations of the data and prior ideas about the relationship between awareness and advertising.

The first consideration was data. The intention was to keep the data costs low and to collect only a minimal amount of data. The authors planned to measure two variables, total brand awareness and advertising

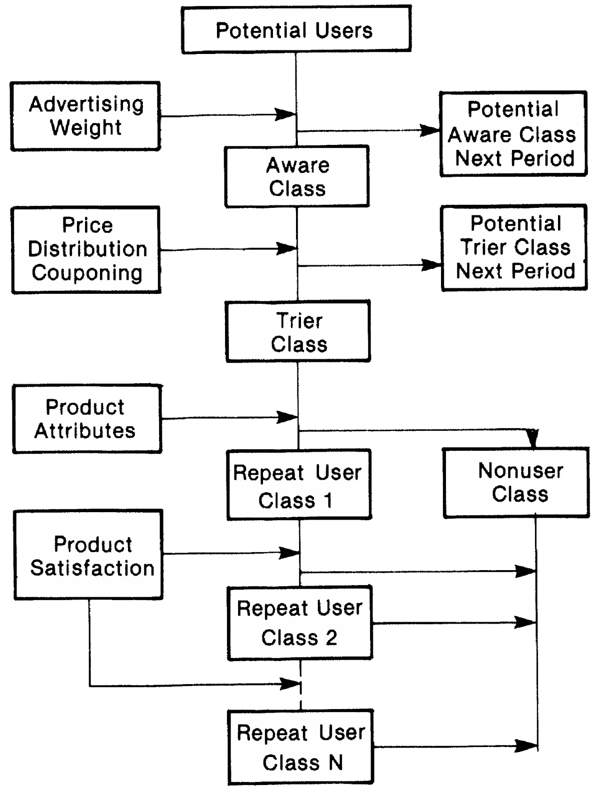


Figure 16–1. Flow Chart of the Model

weight defined in gross rating points (GRP's)[4] Total brand awareness was the dependent variable and the GRP's the independent variable.

Total brand awareness was measured by aggregating the positive responses to two questions among product category users: "What brands of (product category) can you think of?" (a positive response is the mention of the new brand) and, among those who did not respond positively to the first question, "Have you ever heard of (new brand)?" (a positive response would be "yes"). Total brand awareness is the sum of the positive responses to these two questions divided by the total number of product category users.

Many other variables also affect awareness, such as word-of-mouth communications. In theory these variables should be included in the model. If they are omitted and the variables are correlated with the advertising levels, the estimated coefficients will be biased. However, the costs of

Table 16-2
Information Available from Questionnaire

1. Product use	7. Other brands used
2. Frequency of product use	8. Frequency of purchase for each brand used
3. Unaided awareness of brand 1st mentioned 2d mentioned Others	9. Rating of each brand 10. Receipt of free sample 11. Use of coupon
4. Aided awareness of brand	12. Repurchase intentions 1st choice 2d choice Others
5. Length of time since last purchase	
6. Brand bought last	

measuring these other variables would be prohibitive. If, for a given new product, the percentage change of these omitted variables does not greatly vary from period to period, then by using the ratio of this period's awareness to last period's awareness (A_t/A_{t-1}) as the dependent variable, one would not have to worry as much about excluding these other variables from the model. Their effect could be absorbed into the error term without greatly biasing the estimates. Therefore, a function of (A_t/A_{t-1}) was chosen as the dependent variable.

A second consideration in determining the functional form was to design the model to permit diminishing returns to advertising (GRP's). As awareness increases it becomes harder and harder to reach nonaware consumers with media. Further, most of the literature on the effects of advertising seems to indicate that there are diminishing returns to advertising. The mathematical form of the dependent variable in equation 16.1 shows diminishing returns to advertising.

A third consideration is that the model should have the capability to allow awareness to decline if the advertising weight is below some level in a period. In most new product introductions consumers begin losing their awareness of the product unless advertising continually repeats the new brand's name. Thus, if advertising weight is below a threshold frequency, awareness may decline from the previous period's level. Through a parameter of the model, either decreasing or increasing awareness can result for low advertising spending levels, depending on the sign of the parameter.

The mathematical relationship decided on which meets the foregoing requirements is

$$\ln \frac{1 - A_t}{1 - A_{t-1}} = \alpha - \beta\, GRP_t \qquad (16.1)$$

where A_t is the brand's total awareness in period t, GRP_t is the number of GRPs delivered for the brand in period t, and α and β are parameters of the model.

An alternative depedent variable that could have been used is $(1 - A_{t-1})/(1 - A_t)$. This ratio also results in diminishing returns to advertising. However, there are some advantages to using the logarithmic form. If awareness is very high, say .9, to increase it any further would require a very high level of media weight. The logarithmic form of the dependent variable makes it relatively harder to increase awareness by 1 percent at 90 percent than does the nonlogarithmic form.[5] For low awareness, say .1, it is relatively easier for the nonlogarithmic form. It is very difficult to increase awareness when awareness is high because some people have media habits which make it difficult to reach them given limited media funds. Using the logarithmic form of the dependent variable more adequately models this concept.

There are two parameters in equation 16.2: α represents the effect no advertising has on awareness. If α is positive, then awareness declines when there is no advertising. If α is negative, awareness increases when there is no advertising. For any given product category, it may be possible to specify the direction of α before estimation. However, in general, it is assumed that $-\infty < \alpha < \infty$.

And β represents the responsiveness of awareness to advertising. It is assumed that $\beta > 0$ indicating that advertising increases awareness. The larger the value of β, the greater is the effect of advertising on awareness.

A third parameter also must be set, initial awareness. Equation 16.2 is undefined for period 1 unless A_0, initial awareness, is available. Through the use of the econometric methods A_0 can be estimated and it therefore is considered a parameter of the model.

The next step is to fit data to the model described in equation 16.2.

Estimation Results

The estimation results are given for two considerably different product categories.[6] Product category 1 primarily contains new introductions for a well-known company. Consequently, initial awareness was expected to be high due to a bias from consumers stating they are aware of a new product for brands of this company when they are not. Some decay of awareness also was expected if advertising were not very high. However, it was difficult to assess whether the decay rate would be very large. Finally, the category was expected not to be very responsive to advertising. Thus, β would not be very large.

For product category 2, many brands are sold (more than 50), none of which has a very large market share. Initial awareness was expected to be low because there were no dominant brands nor is the company name used in the brand's name (e.g., Del Monte X). For the decay rate parameter a positive increase in awareness was believed possible if advertising levels were low because the product is purchased frequently and therefore consumers could become aware through in-store exposure rather than advertising. However, the authors were not very certain about this outcome or about the effect advertising would have on awareness. Thus, the prior estimates of the parameter's distribution were fairly diffuse.

The results of estimating the parameters are listed in table 16–3. The actual and fitted observations along with the residuals are given in table 16–4. The data given in table 16–4 are for eight introductions for product category 1 and 10 introductions for product category 2. No more than three periods of data are available because the surveying process ends after three waves, In some cases, there is an omitted observation. For example, for product category 1, product 1, period 2 is missing because no survey was taken in that period. Some of the data used for a product category were collected prior to the development of the model. In these cases missing periods were common. However, these data can also be used in estimating the parameters.[7]

For product category 1 initial awareness is very high, .38, and α is positived; β = .0004, which is very low. From the values of α and β one can find the number of GRP's required before awareness does not decline; 385 GRP's per month are needed.[8] This is a fairly high level. Thus, the decay rate is large.

For product category 2, the results are very different. First, initial awareness is low, .16. Next, β = .0007543, which is twice as large as for product category 1; α is also positive but very small. By computing the number of GRP's before awareness declines, one finds that only 4 GRP's are required. Thus, there is very little decay. However, as there are only three periods of data, the decay rate may only represent what happens over the first three periods. In later periods awareness may increase or decline if the GRP level is low, but the change cannot be determined from the available data.

Looking at table 16–4, one sees that the model seems to predict actual awareness fairly well. There are a few cases in which the errors are large. The adjusted R^2's are reasonably high for a time-series cross-sectional model. The average absolute error for product 1 is 7.8 and for product 2 is 8.5. On the basis of both the fits and the reasonableness of the parameter estimates, the model appears to represent the advertising-awareness process fairly well.

Table 16-3
Parameter Estimates for Advertising-Awareness Model

Product Category	A_0	α	β	R^2
1	.38	.0154	.0003995	.732
		(.00905)[a]	(.0000665)	
2	.16	.0031	.0007543	.654
		(.00930)	(.000270)	

[a]The standard errors are the values given in parentheses.

From the values given in table 16-3, it seems that the parameter estimates are very different for the two product categories. If the parameters are not estimated separately but instead the two data sets are pooled and a single set of parameters is estimated, the fitted predictions would be much poorer. This conclusion is based on a casual study of the parameters. However, given the different magnitudes of the parameters (initial awareness is .38 versus .10, decay rate 385 versus 4 GRP's, etc.), it seems necessary to estimate parameters separately for individual product categories. This result is somewhat counter to that of other authors [e.g., see 3]. More theoretical work, beyond the scope of this study, is necessary to explain why product categories should differ in their response to advertising.

Predicting awareness was accomplished by use of only one variable and thus the data costs were low. Because the models seem to fit the data fairly well, the cost of collecting information on additional variables to improve the fit did not seem justified. Therefore the simple model described was used to represent the advertising-awareness relationship.

The Trial Model

The Model

After consumers become aware of the new product, they decide whether or not they should try it. Trial theoretically follows awareness, and therefore triers should be a percentage of the customers who are aware of the new product. However, there are different groups of consumers who are aware and have not tried the brand. One consists of consumers who become aware in the present period ($A_t - A_{t-1}$). The other group consists of those who were aware in the previous periods but have not yet tried the product ($A_{t-1} - T_{t-1}$). The two potential trier groups presumably have different

Table 16-4
Awareness Model's Predictions

Brand Number	*Period*	*Actual*	*Predicted*	*Absolute Error*
Product category 1				
1	1	65	64	1
1[a]	3	91	91	0
2	1	32	48	16
2	2	43	42	1
2	3	48	55	7
3	1	40	38	2
3	2	46	40	6
3	3	38	48	10
4	1	69	60	9
4	2	85	78	7
4	3	84	86	2
5	1	23	42	19
5	3	53	36	17
6	1	49	57	8
6	2	56	57	1
7	1	53	44	9
7	2	53	51	2
7	3	53	70	17
8	1	55	43	12
8	2	69	58	11
8	3	58	65	7
				Average absolute error 7.8
Product Category 2				
1	2	23	7	16
1	3	32	31	1
2	1	30	20	10
2	2	45	45	0
3	2	45	50	5
4	3	58	56	2
5	3	30	38	8
6	2	69	56	13
6	3	76	68	8
7	1	29	47	18
7	2	43	41	2
7	3	32	41	9
8	2	14	41	27
9	3	64	56	8
10	1	24	26	2
10	2	23	34	11
10	3	22	27	5
				Average absolute error 8.5

[a]Missing observations are due to no awareness data available for a product in a given period.

probabilities of trial. In most cases, the newly aware potential triers should have a higher probability of trial than those who have been aware in the past and have not yet tried the new product. These two groups form the pool of individuals available for trial. Thus the model is

$$T_t = T_{t-1} + \alpha(A_t - A_{t-1}) + \beta(A_{t-1} - T_{t-1}) \tag{16.2}$$

where T_t is the cumulative percentage of triers of the new product in period t, A_t is the cumulative percentage of potential triers of the new product who are aware of the new product in period t, and α and β represent the fraction of potential triers who actually become triers in a given period.

The parameters are expected to be constrained by $0 < \beta < \alpha < 1$.

Other factors also should be considered in the model. One is price. Price does not affect awareness, but it does affect trial. Consumers may be interested in trying a new product on the basis of attitudes developed from the advertising, but may not actually try once they learn that the product is twice as expensive as their present brand. Consequently, awareness may be high but trial low because the price is too high. The opposite effect also may occur; that is, low price may induce trial. Couponing is an example of creating high trial by reducing the price of the initial purchase.

Price can be included in the model in several ways. First, one can simply make it an additive term, i.e., $T_t = T_{t-1} + \delta P_t + \ldots$. Another approach, which is the one used, is to adjust incremental trial ΔT_t by a relative price term. The advantage of this form of the dependent variable of the model is that it adjusts trial by a percentage rather than simply reducing trial by some fixed amount as the additive approach does. Using a multiplicative approach also precludes a negative prediction for the change in trial which can happen for the additive model. The trial model now becomes

$$\Delta T_t^* = \alpha(A_t - A_{t-1}) + \beta(A_{t-1} - T_{t-1}) \tag{16.2a}$$

$$\Delta T_t^* = \Delta T_t \cdot \bar{P}_t^{\gamma} \tag{16.2b}$$

$$\Delta T_t^* = T_t^* - T_{t-1}^* \tag{16.2c}$$

$$\Delta T_t = T_t - T_{t-1}$$

where $\bar{P}_t$ is the relative price of the brand at time t and is defined by $\bar{P}_t = P_{b,t}/P_{a,t}$ with $P_{a,t}$ = average price of the product category at time t and $P_{b,t}$ = price of the new brand at time t. If $\bar{P}_t$ is greater than one, then $P_{b,t} > P_{a,t}$ and expected trial is reduced. If $\bar{P}_t$ is less than one, expected trial is increased; $\bar{P}_t$ is raised to the γ power to allow greater or less effect of the price ratio than the simple form $\hat{P}_t$. If $\gamma > 1$, the effect of price differences

across products will be accentuated. If $\gamma < 1$, the difference will be less pronounced.

Another factor that affects trial is distribution. Unfortunately, the available distribution data were extremely poor.[9] Because the model is estimated for each product category and because many of the new products in a specific category are marketed by the same firm, distribution will be approximately the same for many of the brands in a specific product category. Therefore, it may not cause serious estimation problems if distribution is not included in the model.[10]

Certain other factors also affect trial but cannot be included explicitly in the model because of the costs of measuring them. Examples of these factors are word-of-mouth communication, package design, and quality of the advertising message. These factors can be included in the error term of the model, but to do so may cause the error terms for a given brand to be correlated over time. Thus, the error term for period t is $u_t = \rho u_{t-1} + \epsilon_t$ where u_{t-1} is the previous period's error term and ϵ_t is an error unique to period t. Assume the ϵ_t's are uncorrelated over time and across products. The u_t's will be correlated over time. Assume $0 < \rho < 1$, which implies that if last period's error is positive, this period's error will have a higher chance of being positive, and similarly if last period's error is negative, this period's has a higher chance of being negative. The effect of designing the error term in this manner is to adjust implicitly for factors that are not measured but belong in the model.

The magnitude of the error term also should be affected by the price adjustment. Therefore, the variance of the error will be a function of $\bar{P}_t^\gamma$. By simply letting $u_t^* = u_t\bar{P}_t^\gamma$, this problem can be overcome.

The final model becomes

$$\Delta T_{i,t}^* = \alpha(A_{i,t} - A_{i,t-1}) + \beta(A_{i,t-1} - T_{i,t-1}) + u_{i,t}^* \tag{16.3a}$$

$$\Delta T_{i,t}^* = \Delta T_{i,t}\,\bar{P}_t^\gamma \qquad t = 1, \ldots, n_i;\ i = 1, \ldots, k \tag{16.3b}$$

$$u_{i,t} = \rho u_{i,t-1} + \epsilon_{i,t} \quad t = 2,3, \ldots, n_i \tag{16.3c}$$

$$u_{i,t}^* = u_{i,t}\,\bar{P}_t^\gamma \qquad t = 1, \ldots, n_i \tag{16.3d}$$

where $\Delta T_{i,t}^*$ = price adjusted incremental trial for the ith product in period t, ΔT_t = incremental trial for the ith period, $A_{i,t}$ = cumulative awareness for the ith product in period t, $P_{i,t}$ = the relative price in period t for brand i, and $u_{i,t}$ is the disturbance term which follows the process given in equation 16.3c, $\epsilon_{i,t}$ is assumed to have zero mean and to be uncorrelated across time and products, i is the brand subscript, t is the time subscript, n_i is the number of periods of data available for brand i, and k is the total number of brands for the category.

Table 16-5
Parameter Estimates for Awareness—Trial Model

Product Category	*Parameters*[a]			
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\rho}$
1	.274	.048	.2	.1
2	.277	.074	—[b]	—[c]

[a]Standard errors are not reported because nonlinear estimation was used and therefore the standard errors are only large sample approximations.

[b]The prices for all products in category 2 were the same.

[c]Because there were very few products in category 2 for which time-series observations were available, the serial correlation coefficient was not calculated.

Estimation Results

Estimates were computed for the same two products as were used with the awareness model. Table 16-5 gives the parameter estimates, and table 16-6 the fitted predictions and average absolute errors.[11]

The parameter estimates seem to make sense intuitively. First, α and β meet the assumed restrictions that $\alpha > \beta > 0$. For product category 1, $\alpha = .274$ and $\beta = .048$; that is, 27.4% of consumers who became aware in the recent period will purchase in that period, and 4.8% of those who were previously aware but did not try the product will try the brand in the present period. For product category 2, $\alpha = .277$ and $\beta = .074$, which indicates a higher percentage of previously aware consumers who did not try will try the brand in the present period than for product category 1.

For category 1 it was possible to calculate γ and ρ, but for category 2 the price was the same for all brands. Therefore γ could not be calculated. Also because of data limitations, very few time series observations were available, which meant ρ could not be estimated accurately. Only for category 1 are ρ and γ estimated.

The estimate of γ for product category 1 was .23, which indicates the effect of price for this category is less than if the relative price had been used. The brands in category 1 are heavily advertised, and there is substantial product differentiation, which may explain the lack of price sensitivity as indicated by $\hat{\gamma} = .23$.

The estimate of ρ is .1, which indicates a low correlation between the period-to-period error for a specific brand. In fact, setting $\rho = 0$ does not greatly affect the results.

The prediction accuracy of the model seems satisfactory. In only a few cases were the errors large for either product category. The average error

Table 16-6
Trial Model's Predictions

Brand Number	*Period*	*Actual*	*Predicted*	*Absolute Error*
Product category 1				
1	1	21	15	6
1[a]	3	29	31	2
2	1	9	8	1
2	2	7	13	6
2	3	9	10	1
3	1	7	10	3
3	2	8	10	2
3	3	9	10	1
4	3	28	27	1
5	1	5	6	1
5	3	23	14	9
6	1	7	12	5
6	2	10	10	0
7	1	15	15	0
7	2	17	17	0
7	3	22	19	3
8	1	11	14	3
8	3	16	16	0
				Average absolute error 2.4
Product category 2				
1	2	11	10	1
2	1	12	10	2
2	2	20	17	3
3	2	15	14	1
4	3	27	19	8
5	3	8	11	3
6	2	16	22	6
6	3	23	23	0
7	1	2	8	6
7	2	5	8	3
7	3	−3	2	5
				Average absolute error 3.5

[a]Missing observations are due to no trial or no awareness data for a product in a given period.

was very low, 2.4 and 3.5 for categories 1 and 2, respectively. When the prediction accuracy and the signs and magnitudes of the coefficient estimates are combined, this model appears to be very good.

One last point deserves discussion. Because only three months of data are used, all past aware but nontriers were aggregated into a single variable, $A_{t-1} - T_{t-1}$. However, the longer the time that a consumer is aware and does not try, the less likely he is to try eventually. In use of the model as a

planning tool, it may be sensible to adjust β downward after some time period.[12]

The Projection Model

The final stage is the projection model, the goal of which is to forecast year-end market shares (or sales) of the new product using as inputs forecasts of trial rates made from the trial model, repeat usage rates, repeat purchase proportions, and trial usage rates.

The idea underlying the projection model is simple. The consumer begins as a trier. He can then either use the product again, becoming a repeat user, or enter the nonuser class. He then either continues to use the product or enters the nonuser class.[13] This process continues each period after initial trial with the consumer either staying in the user class or moving into the nonuser class. The model does the tracking of the percentage of consumers in each stage of repeat usage. To compute each period's sales, the model simply aggregates the percentage of users in each repeat class and the period's new triers. Figure 16–2 is a flow chart of the projection model.

As seen from figure 16–2, to make year-end market projections, one needs trial rates, the percentage of repeat user, consumption rates for triers and repeat users, and a decay rate for repeat users. The next section shows how these inputs were combined to project year-end market share or sales.

Projection Model Equations

The basic mathematical relationships used to make projections are:

$$UC_{t-1}(t) = r\,\Delta T(t-1) \qquad t = 2, \dots \qquad (16.4a)$$

$$UC_{t-1}(t) = [1 - d(i-1)]\,UC_{t-i}(t-1) \quad i = 1, \dots, t-1 \qquad (16.4b)$$

where $\Delta T(t)$ = the percentage of new triers for period i

r = the percentage of triers in period t who remain users in period $t+1$

$UC_{t-i}(t)$ = the percentage of new triers in period $t-i$ who are still in the user class at time t

$d(i-1)$ = the percentage of new triers in period t who purchased in period $t+i-1$ and stop purchasing in period $t+i$

t = the time period, which is usually in months

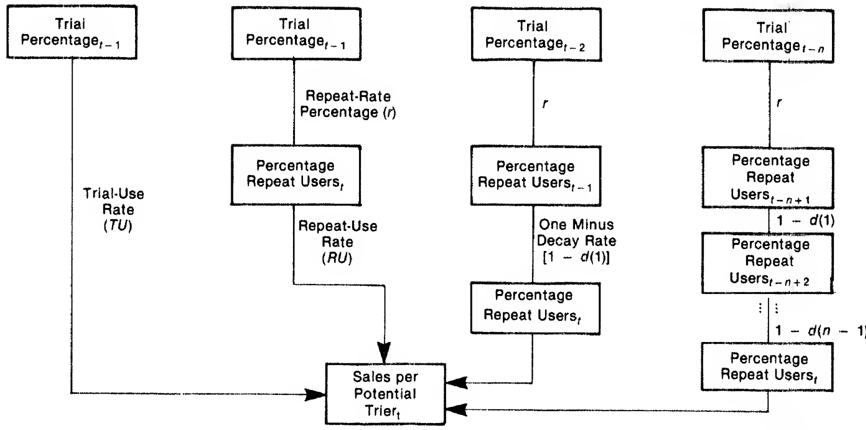


Figure 16–2. Flow Chart of the Projection Model

Equation 16.4a says that r proportion of new triers in period t will repeat purchase in period $t + 1$. Equation 16.4b says that $1 - d(i - 1)$ proportion of users who purchased in period $(t + i - 1)$, will repeat purchase in period $(t + i)$. Thus, equations 16.4a and 16.4b compute the proportion of new triers in period t who repeat purchased in period $t + 1, t + 2, \ldots$.

To compute total sales in a given period, one multiplies the proportion in each user class by their usage rate. Then total sales per potential trier[14] in period t, $TS(t)$, is simply

$$TS(t) = TU\,\Delta T(t) + \sum_{i=1}^{t-1} UC_i(t)\,RU \qquad t = 2, 3, \ldots \qquad (16.5)$$

$$TS(t) = TU\,\Delta T(t) \qquad t = 1$$

where TU = the quantity purchased by a trier, and

RU = the quantity purchased by members of the user class

In equations 16.4a and b and 16.5 there are several unknown parameters, $d(1), \ldots, d(12)$, r, TU, and RU, that must be estimated. To estimate the parameters, some simplifying assumptions are made. To begin, it is assumed that the repeat usage percentage r is not dependent on time. This is obviously not true. However, if r is time dependent, one must estimate 12 parameters. A simplifying assumption is that the repeat usage percentage is constant. Evidence about early adopters versus late adopters contradicts this assumption. However, the authors have found accurate pre-

dictions assuming a constant repeat rate, and therefore as an approximation it may not be too costly, in a prediction sense, to assume the same value of r for all periods.[15]

To estimate the other parameters of the projection model it is also assumed that they are not dependent on time. Note that $d(i)$ still depends on the length of time from the first period the consumer tried the product, but it does not depend on the length of time since the new product was introduced.

To compute year-end sales one can sum the sales for each month, $\Sigma_{t=1}^{12} TS(t)$, and obtain a forecast. Usually firms are interested in longer term sales in addition to first year sales. To forecast future sales, sales for the last few months of the first year are used to make projections. The reason is that new product sales usually spurt in the early months of the introduction and then dampen after four or five months. By months nine or ten, sales have usually reached "steady state." Thus, the last few months of the first year are more representative of future sales than are the total first year sales.[16]

Parameter Estimation

The parameters that need to be estimated are TU, RU, $d(t), \ldots, d(12)$, and r. Two of these parameters, r and RU, were estimated by using telephone surveys done monthly for the first three months of the introduction.[17] The $d(i)$'s were estimated subjectively by using repeat rates and product satisfaction information received from the questionnaire. (The reasons for using subjective estimates are explained hereafter). TU was set equal to one because triers automatically use one unit in the period they become triers.

Two procedures were used to estimate r. In the first period after the introduction, most triers have not had time to repurchase. Consequently, questions about purchase intentions and product satisfaction measures were used to estimate a repeat purchase probability. If the customer seems to have had a positive experience with the product and intends to purchase, he is classified as being a repeat user. If he says he intends to repurchase but then expresses negative opinions about the product, he is classified as a nonrepeat user. The percentage of tries classified as users gives an estimate of r.

For periods two and three, actual repeat rates can be used. Customers who first bought more than one period ago and have repurchased in the last period are classified as repeat users. The percentage of customers who repurchase divided by the total number who have tried the newproduct more than one period ago gives an estimate of r.

The three estimates of r (one for each wave) are averaged to give a final estimate of r. Even though the first period's estimate is subjective, it seems

to be close to the last two in almost all of the new product introductions that have been studied. This averaged estimate of r is used as an input to the projection model.

To estimate the repeat purchaser's usage rate RU, data collected from the telephone survey were used. A repeat user is not required to be "brand loyal" to the product. He is only required to have used the new product in the last period. The repeat usage rate measures the consumption rate for the new brand by repeat users.

The final parameters to be estimated are the decay percentages for each user class $d(1), \ldots, d(12)$. Estimating the $d(i)$'s requires data about when the product has been first purchased and the repurchase rates for users who tried one, two, three, . . . periods ago. In the first three months of the introduction, one usually gets information about trial and one repurchase. Occasionally there are two or three repurchases. However, only a few, if any, triers have repurchased often enough for estimation of $d(1)$. Almost never can one estimate $d(2), \ldots, d(12)$. Therefore, the authors decided to use subjective estimates of the $d(i)$'s.

To obtain subjective estimates for $d(1)$, assume that the probability of repurchasing given trial and one repurchase is higher than that of simple repurchasing after initial trial. Thus $r < [1 - d(1)]$. Using a similar argument, assume that $[1 - d(1)] \leq [1 - d(2)] \leq, \ldots, \leq [1 - d(12)]$. From information about purchase intentions and product satisfaction among repeat users, one can only determine whether repeat customers seem satisfied. If there seems to be dissatisfaction, then the estimates of the $[1 - d(i)]$'s are lowered.

Because subjective estimates of the $d(i)$'s were used, a sensitivity analysis of the $d(i)$'s was applied to show whether the resulting forecasts change drastically. An example of the sensitivity analysis is given in table 16-7. To show the effect of different values of $d(i)$'s, the same values of r, RU, and $\Delta T(t)$ were used for all three runs; r was set at .4, RU was set at 1.4 units per period, and the trial figures are listed in table 16-7. The output of the projection model represents unit purchased per potential customer. For example, in table 16-7, for all three sets of $d(i)$'s, .08 unit per potential trier was consumed in period 1. In other words, each potential consumer used eight hundredths of a unit in period one. The total number of potential consumers for this product category would be multiplied by .08 to obtain a sales estimate for period one.

The results show that year-end sales for the first set of (low decay) $d(i)$'s are approximately 33% larger than those for the third set of $d(i)$'s (high decay). This finding indicates that the sales results are somewhat sensitive to the values of the $d(i)$'s used. However, it is important to note that even though each $d(i)$ may not look very different between set one and set three, their aggregate effects are very different. For a trier with $d(1) = .2$

Table 16–7
Sales per Potential Trier for Different Decay Percentages

		Repeat-Rate Percentage = 0.4, Repeat-Usage Rate = 1.4											
Period		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Trial percentage		8	5	4	2	2	1	1	1	1	.5	.5	.5
		Decay Percentage [d(i)]											
Period		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Decay rate	Low	20	10	5	4	3	2	2	2	2	2	2	2
	Medium	25	15	8	5	3	2	2	2	2	2	2	2
	High	30	20	10	5	3	2	2	2	2	2	2	2
		Sales per Potential Trier											
Period		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Decay rate	Low	.0800	.0948	.1038	.0971	.0999	.0948	.0948	.0963	.0983	.0954	.0946	.0947
	Medium	.0800	.0948	.1016	.0920	.0921	.0853	.0841	.0849	.0864	.0831	.0822	.0819
	High	.0800	.0948	.0994	.0871	.0851	.0771	.0752	.0753	.0764	.0729	.0716	.0712

and $d(2) = .1$, there is a 29% chance of his using the product three periods later ($.4 \times .8 \times .9$). With $d(1) = .3$ and $d(2) = .2$, he has a 22% chance of using the product three periods later. Therefore, there is a considerable difference in the two sets of $d(i)$'s used.

To use the projection model, one must input one more parameter, the trial usage rate *TU*. This parameter was set equal to one because a trier, by definition, must purchase one unit in the period in which he initially tries the product.

From the description just given, one sees that two parameters of the projection model are estimated from each period's telephone survey, *RU* and *r*, one set of parameters is estimated subjectively, $d(1), \ldots, d(12)$, and *TU* is set equal to one. As one obtains more experience with the model, it becomes easier to estimate the $d(i)$'s. If consumer diary panel data were used, more accurate estimates of these parameters would be obtained. However, on the basis of the forecasting accuracy of the model with the estimation procedures described, the additional cost of panel data does not seem justified.

Using the Projection Model When the New Product Is Sampled Heavily

When a new product is sampled (free samples are given to potential users), the awareness-trial-repeat usage chain is broken. Trial occurs simultaneously with awareness. The trial rates may be different for consumers given samples than for consumers who must pay for the product in order to try it. If a consumer has tried the product but did not purchase it, the reason may be that it is more expensive than he had expected. Therefore, a sampled customer's repeat rate will probably be different from that of a nonsampled customer. Other factors also may intervene which will make his repeat rates different from those of nonsampled customers. Therefore, in estimating the parameters of the projection model, the sampled and nonsampled triers are separated. Usually sampled customers have both higher trial rate (more like initial repeat usage rate for nonsampled consumers) and higher repeat usage rates. By knowing the initial percentage sample, one can easily combine the sampled and nonsampled groups to make year-end sales projections.

Diagnostic Information from the Projection Model

The projection model is also used to diagnose problems with the introduction. Because the main factor in the ultimate success of a new product is

repeat purchasing, it is useful to determine why a product has a low or high repeat rate. This is done by asking questions about product satisfaction and repeat purchase intentions. From these questions one often can learn why a new product is failing or why it is successful. The latter is important because if a product has a high repeat user percentage but appears to have relatively low product satisfaction, repeat rates may soon begin to decline. Thus the decay percentage will be higher than normal.

If a product has a low repeat usage percentage, one can often identify the causes. Perhaps there are unexpected design problems, or the product is too similar to other products on the market. From the diagnostic information, advice can be given about how to redesign the product so that it can be introduced in other test markets and retested.

Obviously, the diagnostic feature is an essential component of the projection model. Models using panel data such as the Parfitt-Collins model [11] would require additional research because diagnostic information is not a natural byproduct of the model. One can be more confident in making projections when repeat-usage percentages correspond to consumer statements about product satisfaction and purchase intentions.

The Planning Model

The model described in the preceding sections also can be used as a planning model prior to test marketing. In many cases firms are interested in evaluating how several marketing strategies affect sales. By specifying a media spending strategy, price, and sampling level plus some estimate of usage and repeat rates, the model can make pre-test market sales forecasts. These forecasts are based on "norms" for the product category. Often products exceed or fall below the norms.[18] However, using the planning model gives the firm some indication of the effects of alternative strategies.

The most difficult inputs to obtain for the planning model are the repeat usage percentage and trial and repeat usage rates. Because these parameters are extremely important, usually a number of alternative rates are used. The normal repeat usage percentage is about 40%. Successful products often have values around the 50% level and unsuccessful products around 30%. One can make projections using all three rates to see their effect on sales.[19] In addition results from concept tests and in-home use tests are often helpful in making pre-test market estimates.

Using the planning model has helped firms determine the potential magnitude of sales expected from the test market. In one study, the firm's prior sales estimate was 6 million case sales for the first year. Use of the planning model with optimistic repeat usage percentages and high trial rates generated only 2 million case sales. Therefore, the firm was able to revise its initial production requirements downward to a more appropriate level.

In another case a firm evaluated different advertising budgeting strategies to see what effect they had on sales. Given the cost of goods sold, the firm could do a profitability analysis to determine which strategy led to maximum profits.

Conclusion

A new product forecasting model that is fairly accurate and inexpensive to use is presented. The model uses survey data rather than panel data. It also uses marketing inputs such as advertising and price. The model can forecast year-end sales after only three months of test marketing so that a quick decision about the success of the product can be reached. The model also gives diagnostic information about the reason for low awareness, trial, or repeat purchasing. It thus helps firms to redesign the product or the introduction strategy. Finally, the model also can be used for evaluating alternative marketing strategies. The next step is to develop accurate inputs to the model before test marketing so that the model can be used to predict success or failure of products without requiring test marketing. Thus the high cost of test marketing can be avoided.

Notes

1. The forecasting and planning models are being used by Leo Burnett Co. The results given are for forecasts that they have made by using the model.
2. Deviations in timing of these waves occur when the purchase cycle of a product is exceptionally long or short.
3. For product categories for which available data are not sufficient for estimation, general norms are used or in certain situations norms for similar products are used (e.g., cat food for dog food).
4. GRP's are equal to the percentage of TV households reached at least once with the advertised message multiplied by the average number of times the household is exposed to the message.
5. "Relatively harder" means that the ratio of GRP's required to increase awareness 1% at 90% is higher than for the nonlogarithmic model. Of course, one must make the model equivalent at some other point to make this comparison.
6. The procedures used to estimate the model are available from R. Blattberg.
7. Some economic adjustments were necessary to estimate parameters for categories missing observations. For example, weighted least squares was used to adjust for the change in the variances of the errors. Because of

the complexity of the adjustment procedure and because new data do not contain missing observations, the exact estimation steps are not given.

8. The high number of GRP's to maintain awareness is due to the nature of the product category. There are a large number of brands and heavy advertising.

9. The problem with distribution estimates is that they vary depending on the data source. The source of the distribution data varied within a category, and therefore these data were not used.

10. Distribution statistics are still used for diagnostic purposes.

11. The estimation procedure used for the awareness to trial model is available from R. Blattberg.

12. Parfitt and Collins [11, p. 136] state that late triers have lower average repeat purchasing rates.

13. Aaker [1, p. 448] in his new trier model found that consumers replace their old brand with the new brand or they discard the new brand if they do not like its attributes.

14. To compute total sales, the number of potential triers (users) of the product class is multiplied by $TS(t)$.

15. Note that in many models parameters probably vary from period to period. As an approximation they are often assumed to be fixed.

16. If the product is seasonal, then a seasonal adjustment needs to be made before using the last three months.

17. Occasionally shorter or longer time periods are used because of the length of the purchase cycle of the product.

18. Forecast accuracy for the model will be less than that given in table 16-1 when the planning model is used because test market data are not used for the first three periods.

19. To facilitate these calculations, a time-sharing version of the model has been developed which allows the model's users to test alternative strategies very rapidly.

References

1. Aaker, D.A. "The New Trier Stochastic Model of Brand Choice," *Management Science,* 17 (April 1971), B435-50.
2. Ahl, D.H. "New Product Forecasting Using Consumer Panels," *Journal of Marketing Research,* 7 (May 1970), 160-7.
3. Claycamp, H.J., and L.E. Liddy. "Prediction of New Product Performance: An Analytical Approach," *Journal of Marketing Research,* 6 (November 1969), 414-20.
4. Eskin, G.J. "Dynamic Forecasts of New Product Demand Using a Depth Repeat Model," *Journal of Marketing Research,* 10 (May 1973), 115-29.

5. Fourt, L.A., and J.W. Woodlock, "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* 24 (October 1960), 31-8.
6. Johnston, J. *Econometric Methods.* New York: McGraw-Hill Book Company, Inc., 1972.
7. Learner, D.B. "Profit Maximization through New Product Marketing Planning and Control," *in* F.M. Bass, C.W. King, and E.A. Pessemier, eds., *Applications of the Sciences in Marketing Management.* New York: John Wiley and Sons, Inc., 1968, 151-67.
8. Massy, W.F. "Stochastic Models for Monitoring New Product Introduction," *in* F.M. Bass, C.W. King, and E.A. Pessemier, eds., *Application of the Sciences in Marketing Management,* New York: John Wiley and Sons, Inc., 1968.
9. NEWS: *N*ew *E*arly *W*arning *S*ystem, Technical Report, Management Science Department, BBDO (no data given).
10. NEWS. The NEWS Model, A Technical Description, Technical Report, Management Science Department, BBDO.
11. Parfitt, J.H., and B.J.K. Collins, "Use of Consumer Panels for Brand-Share Prediction," *Journal of Marketing Research,* 5 (May 1968), 141-45.
12. Silk, A.J., and G.L. Urban. "Pre-Test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," MIT working paper, 834-76, January 1976.
13. Urban, G.L. "Sprinter Mod. III: A Model for the Analysis of New Frequently Purchased Consumer Products," *Operations Research,* 17 (September-October 1969), 805-54.
14. Urban, G.L., and Richard Karash. "Evolutionary Model Building," *Journal of Marketing Research* 8 (February 1971), 62-71.

17 Application and Utilization of Test-Market-Based New-Product Forecasting Models

Joe A. Dodson

All too often, test markets are used like a thermometer to test the temperature of the market for a particular product. New-product managers often use the results to look at a single number—market share. This is a terribly expensive and inefficient use of the resources required to put a new product into test market. All the new-product manager can do is carry his market-share number back to his office and start scratching his head to find ways he can improve his new product's performance. This often means finding ways to explain away the product's below-expectations performance. The use of test-market-based new-product forecasting models as an alternative to head scratching is the focus of this chapter.

Objectives of Test Markets

The objective of a test market is to evaluate consumer reaction to a new product under normal marketing conditions. The primary motive for testing is to obtain an estimate of potential product sales. A secondary motive is to pretest alternative marketing plans. Used correctly, test markets provide a wealth of information for interpreting the new product's performance. And, new-product models are being used increasingly by innovative companies to interpret test-market information and to predict the commercial viability of new products.

Value of Test-Market-Based New-Product Forecasting Models

The introduction of a new product is a complex and difficult task. There are important interactions among distribution levels, market-category development, display factors, and sales-force strength that will vary from market to market and must be taken into account before a new-product manager can derive a reasonable estimate of national sales.

Some companies are using two, three, or more test markets at a time to try alternative introductory marketing strategies. All these data can be useful in understanding consumer response to the new product if they are used properly. These data will be most useful if they are combined and interpreted against a consistent view of the world. Only then can the relative importance of each element of the introductory strategy be accurately evaluated and important interaction effects detected. In my experience, test-market-based new-product forecasting models provide such a structured filter in which variables can be quantified, their relative importance can be evaluated, and consistent comparisons can be made.

Test-market data are of little use if they do not explain why the new product obtained a specific level of market share. Test-market results can be used to provide precise parameterization of a manager's own judgment about the relative importance of each of the factors which will determine the market's response to the new product. Used properly, they are a significant aid to decision making. They open up avenues for the new-product manager to explore and identify the weak points in his new-product strategy.

In the final analysis, the most valuable model is the one which enables the new-product manager to evaluate and forecast the results of alternative courses of action. The influence of packaging, pricing, advertising, positioning, and promotion must be understood if a model is to be expected to help the new-product manager make better decisions.

One of the first new-product models which dealt with the impact of controllable marketing-mix variables was developed at N.W. Ayer (Claycamp and Liddy 1969). It has evolved over time, and its current application at N.W. Ayer is described in more detail below.

I am not aware of any conclusive scientific evidence that new-product decisions have been improved with the use of these models. But I am aware of, and have personally experienced, direct benefits through the application and utilization of the Ayer new-product model with several major consumer packaged-goods clients. This chapter focuses primarily on how test-market-based new-product forecasting models are being used today and the opportunities and problems they present to the new-product manager.

Experience with Test-Market-Based New-Product Forecasting Models

Most major packaged-goods companies today have a new-products group. This group usually consists of a vice president and one or more new-product managers. Some formalized new-product development process is usually

adopted and espoused by the group. And test marketing is virtually always a part of this process.

Test marketing is usually a critical stage in the new-product development process. But test-market-based new-product models are seldom used to understand and evaluate test-market results. In most cases this is because new-product managers do not understand the value of models to interpret test-market results. A good model will force explicit consideration of the dimensions of consumer response which affect a new product's performance. Focused evaluation on each of the elements which determine first-year sales can help pinpoint information.

The need for this attention to detail becomes clear when we recognize that all too often first-year sales goals are set arbitrarily by new-product managers at 50,000 units, 100,000 units, or some other suitably round number without any consideration of the assumptions which must be met before the new product can attain these goals nationally. Test markets and test-market-based new-product forecasting models force explicit consideration of the assumptions being made about the component parts of the process which generate sales volume and provide more reasonable and realistic sales goals.

I have also watched companies launch new products nationally assuming a high market share based on early test-market performance, only to find out too late that the high share was due to unexpectedly high trial. And a low repeat rate followed which could not sustain the product nationally. But because there was no effort to track the component parts of the new product's performance, this was not known until after major investments had been made in the national launch of the product. Utilization of test-market-based new-product models which incorporate the dynamics of consumer trial (Eskin 1973; Parfitt and Collins 1968; and Fourt and Woodlock 1960) are being used increasingly by companies to reduce the likelihood of such costly mistakes.

Experience suggests that such models should initially be simple. Models can be improved and enhanced over time when the new-product manager is ready for more complexity. Even a simplistic three-variable model like $S = T^*R^*P$, where S = first-year sales, T = number of triers, R = repeat rate, and P = purchase rate, can help avoid simple mistakes. A simple test-market-based new-product model can evolve over time to a more complex and sophisticated model like Urban's (1970) SPRINTER Mod III. But to be useful, such a model requires a major commitment of management's time and company data. There are a considerable number of parameters, many of which may be remote from the usual experience of many new-product managers. Nonetheless, the development and use of a test-market-based new-product model will evolve (Urban and Karash 1971) over time as a company gains experience and understanding.

This discussion points up a dilemma frequently faced by a company trying to introduce and use test-market-based new-product models. The new-product manager wants a "complete" model that is *representative* of his product category. An attempt to achieve completeness usually leads to the development of a complex model. But complexity is the last thing the new-product manager wants. He is looking for the model to simplify his life. Seldom will the new-product manager use a model he does not understand. Even in those cases where the model is a standardized part of the new-product development process, it is often maintained in the corporate research department. This organizational separation often impedes the effective use of the model. The statistician is usually interested in the statistical efficiency of the model's parameters and its predictive accuracy. He often has little understanding of and appreciation for the problems the new-product manager faces.

Both these problems can be overcome. Experience indicates that the value of the model is substantially enhanced if the model is structured in terms the new-product manager can understand and if the variables are defined in operational terms. In my experience, a new-product manager is more willing to use a model if he feels that the model "speaks his language." This problem and many others which arise can also be overcome if there is someone, say a marketing scientist, who can provide an effective interface between the model and the new-product manager. This person must understand and appreciate the problems faced by the new-product manager. Otherwise, he will have difficulty obtaining the manager's confidence in the model. He must also be able to explain the model to the new-product manager. The new-product manager would rather trust his judgment than a black box to understand why his product performed as it did in test market.

When to Use a Test-Market-Based New-Product Model

Test-market-based new-product models were initially developed for the express purpose of forecasting national sales on the basis of test-market results. With test-market results, a model can be parameterized by using actual empirical data. Increased competition has forced many new-product managers to look for ways to expedite the time lag between new-product conception and national launch. This has led to the use of test-market-based forecasting models prior to test market.

Using these models prior to test market has become increasingly popular. Changes can be made to a new product more easily before major resources have been expended on packaging, advertising, or capital expenditures for production to enter test markets. Use of a model prior to test

market helps to identify which variables need to be tracked in test market. Models which are defined in operational terms make it easier and less expensive to monitor the important dimensions of a new product's test-market performance. If these considerations have been ignored prior to test market, it is highly likely that an important variable will not be measured, making it difficult to parameterize a model to predict national performance after test-market results are in.

As noted earlier, new-product managers often go into a test market watching one factor—market share—only to find out too late that the test-market results are not useful for projecting national sales because market conditions changed during the test. A competitor may have entered with a competitive product. Or the raw-material costs may have risen in the last three months, requiring the product manager to raise his price prior to national launch. It is usually too late to try to go back and collect the data needed to understand the "why" of the product's performance after the test-market results are in. Experience indicates that getting a new-product manager involved in the model prior to test market will prove extremely helpful in gaining his commitment for using it to evaluate test-market results.

The AYER New-Product Model

A model which was developed specifically for use prior to test market is the AYER new-product model. An early version of this model was described by Claycamp and Liddy (1969). The model has since been revised and enhanced. And it has been used both before and after test markets to interpret results and predict new products' national sales.

The variables now used in the model are listed in table 17-1. The equations described by Claycamp and Liddy have been replaced with modified logit curves. This led to more complex estimation problems which required the use of nonlinear regression procedures. But the replacement of the linear equations with S-shaped curves has also eliminated the possibility of unreasonable predictions and improved the predictive accuracy of the model. These enhancements have made the original model more robust.

The Ayer new-product model was designed to predict advertising recall and initial trial thirteen weeks after the new product's introduction. Since the primary purpose of a new-product introduction is to stimulate trial, most campaigns are planned in terms of thirteen-week cycles. But the eventual success or failure of consumer packaged goods depends on their ability to generate repeat sales beyond the first quarter.

Since the publication of its early efforts, N.W. Ayer has added additional models to its model bank. These are tied together into a package

Table 17-1
Description of Variables Used in N.W. Ayer New-Product Model

Category interest	Index of consumer interest in the product category
Product positioning	Judged strength of "reason why," product uniqueness
Communication	Judged effectiveness with which product position is conveyed in advertising
Persuasion	Judged conviction of selling proposition
Media impressions	Number of gross household impressions adjusted for probability of delivery
Nonmedia impressions	In-home impressions delivered through promotions
Category noise level	Annualized rate of competitive spending
Advertising recall	Percentage of households able to accurately recall advertising claims at end of thirteen weeks
Monthly purchase rate	Frequency of buying
Distribution	Share of all commodity value in chain stores
Display	Share of category facings plus display
Packaging	Judged visibility, adaptability, and impact of package
Consumer promotion	Score based on type, value, and coverage of consumer promotions
Switchability	Willingness of consumers to try new product in category
Family branding	Known or family brand name relevant to product category

called SPECS (*s*trategic *p*lanning *e*valuation and *c*ontrol *s*ystem) (figure 17-1). SPECS goes beyond a prediction of first-quarter trial. It is a structured, model-based approach to combining the best available information on factors such as consumption rates, product costs, and investment requirements. SPECS includes various cost, cannibalization, and profit submodels which are used to generate estimated profits and return on investment over a company's planning horizon. With these enhancements we are able to calculate the expected rate of return on investment and the probability that a new product will exceed a company's required target rate of return.

The N.W. Ayer new-product model provides a prediction of first-

quarter trial for a new product. This prediction is the key prediction which derives the output from SPECS. The first quarter was chosen as the critical time frame within the model. Experience has indicated that if distribution and shelf take-away are not obtained within the first thirteen weeks in a market, there is little chance for new-product success. The important first-

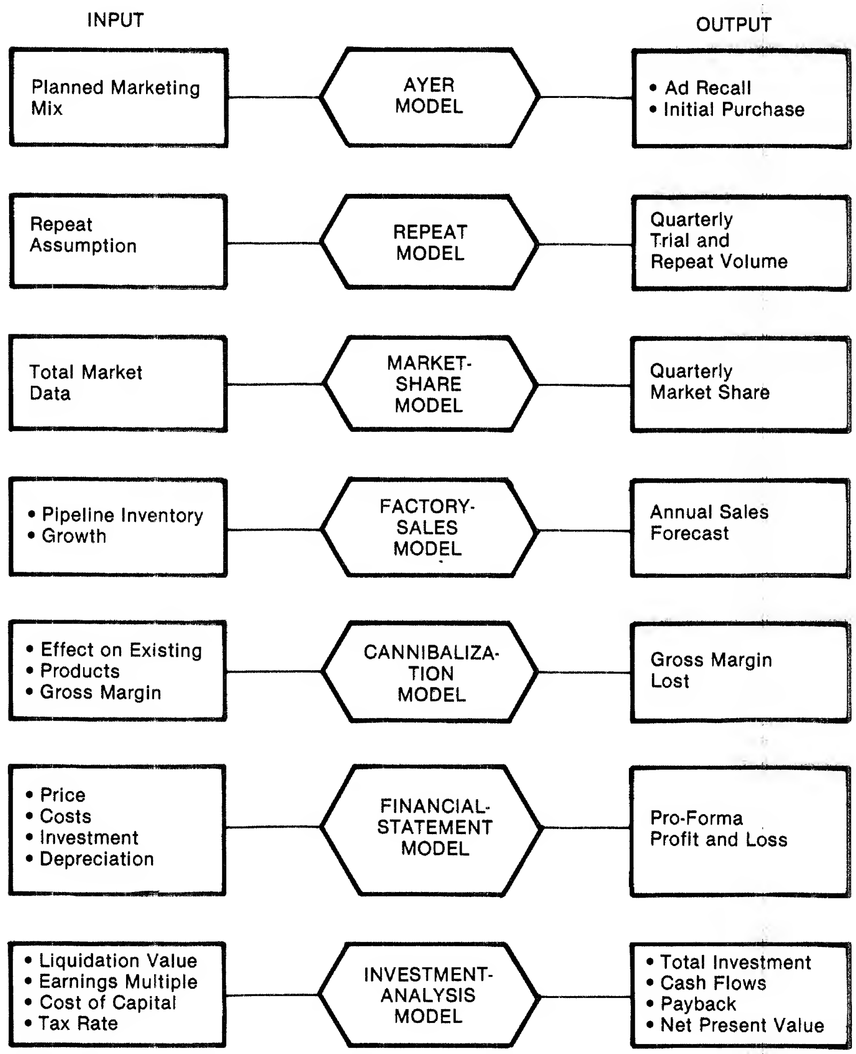


Figure 17–1. N.W. Ayer Strategic Planning Evaluation and Control System

quarter trial estimate is then combined with a "typical" cumulative penetration curve for the product category in the penetration/repeat-purchase submodel to provide an estimate of quarterly trial in the first year. The new-product repeat rate, derived from managerial judgment, in-home use tests, or test-market experience combined with the trial estimates, provides an estimate of total first-year sales volume. The output from each submodel provides the basic input into the next submodel in SPECS.

An important contribution of SPECS to the original Ayer new-product model has been the extension from a model designed only to predict sales to a model which forecasts the product's worth to the company. This has increased the value of the model. And in many cases it has led to its integration into client companies' financial planning. It is all too easy for a new-product manager to be so focused on his product as to lose sight of the overall corporate objectives. SPECS explicitly takes into account category growth, likely competitive retaliation, and expected cannibalization and evaluates their impact on company revenues and profits.

Let me use the Ayer new-product model, with which I am most familiar, to further illustrate some of the ways test-market-based new-product models are being used. The Ayer new-product model is unique in that it substitutes independent judgment for the new-product manager's judgment in determining the input values for factors in the model. Consistency is maintained by using standardized rating procedures and an experienced panel of marketing and advertising professionals. The developers of the model strove for completeness even though this required the use of judgment. This completeness has substantially enhanced the diagnostic value of the model. It gives the new-product manager independent feedback on his product's positioning, packaging, branding strategy, and so on, prior to test market.

The efficacy of these inputs as well as all the factor scores used in the model is ultimately judged by test-market results. However, direct consumer response prior to test market can also be used. For example, concept tests based on quantitative surveys or in-home use tests can provide an index of the relative effectiveness of alternative product positionings. Controlled tests of different package designs can provide a consumer-based score of the packaging factor, which can be used as input into the model. Though not required, consumer research can provide validation of the model's input-factor score.

The model is used to identify the strengths and weaknesses of a new product's introductory marketing strategy by comparison with the over 100 cases now in the Ayer new-product model data bank. This data base contains detailed information tracked by N.W. Ayer on the new-product introductions of over forty major consumer packaged-goods companies, including Proctor and Gamble, General Foods, General Mills, Kraft, Pillsbury,

Nestle, Ralston Purina, and Philip Morris. Additional input information about repeat rates, prices, costs, and so on, enable the new-product manager to use SPECS to obtain pro forma profit-and-loss statements for his brands (figure 17–2). This provides a very early reading on the brand's most likely impact on corporate profitability.

These enhancements to the original model have led to increased utilization of the model. Many companies now have new-product review boards headed by the vice president in charge of new-product development. These boards often require the new-product manager to estimate his new product's impact on corporate profits. SPECS provides the new-product manager with an independent evaluation of the financial impact of his product on the corporation.

The real value in the model is realized when the model is used by the new-product manager to experiment with marketing strategy. Once the model is calibrated, using test-market data, the new-product manager can ask "what if" types of questions without incurring the cost and time delays required by further test markets. Directions for further work and research on the new product are pinpointed by testing the sensitivity of the new product's sales to different model inputs. This enables the new-product manager to concentrate his resources to strengthen the weak spots in the marketing strategy for his new product. As new data become available from in-home use tests, advertising-copy testing, experimental laboratories, or further test markets, the model's inputs can be revised and its predictions updated. The Ayer new-product model, when used to full advantage, provides direction to management for finding the path which most efficiently maximizes expected profit.

Growing Use of Models

In spite of their advantages, the increased use of test-market-based new-product models for evaluating alternative marketing strategies continues to grow slowly. Several factors explain this fact. The collection of test-market data requires planning in advance. Many new-product managers are on "fast tracks" and are rewarded for quick success. This results in a short-sighted use of test-market data. In addition, a major corporate commitment is required to invest the time and money required to build a consistent base of data from prior new-product introductions.

Expanded use of test-market-based new-product models is not likely to come from the development of more complex or more sophisticated models. Rather, progress will come from learning how to build models which are more easily integrated into the company's day-to-day process of new-

Quarter	Sales	Trial	Repeat	Percent Repeat	Percent Share
1	567,836	493,771	74,066	13.0	2.1
2	267,201	11,916	147,285	55.1	1.0
3	378,440	70,539	307,901	81.4	1.4
4	338,586	21,162	317,424	93.7	1.3
	1,552,063	705,387	846,676	54.6	1.4

Pro Forma Profit-and-Loss Statement ($1,000)

	Year 1	Year 2	Year 3	Year 4
Gross Sales (1)	29,178	21,675	22,759	23,442
Net Sales	28,011	20,808	21,848	22,504
Cost of Goods Sold	17,609	13,081	13,735	14,147
Gross Margin	10,402	7,727	8,114	8,537
Development Costs	500	0	0	0
Marketing Costs	5,700	2,600	2,600	2,600
Depreciation	100	100	100	100
Gross Contribution	4,102	5,027	5,414	5,657
Cannibalization	−671	−997	−1,047	−1,078
Net Contribution	3,431	4,030	4,367	4,579

Figure 17–2. First-Year Retail Sales (Cases)

product development. Forced use of a new-product forecasting model often appears to the product manager as an obstacle for *his* new product to overcome. The integration of models into the new-product organization and the infusion of "marketing scientists" who understand marketing *and* models will do more to stimulate the application and utilization of these models.

In spite of the problems that exist, we can expect the use of test-market-based new-product models to continue to grow in the future for several reasons. The costs of investment required to produce and market new products have continued to increase. Advertising and promotion campaigns for national new-product introductions easily approach and often exceed $10 million. New-product managers and corporate management have become increasingly aware of and familiar with the value of models in marketing. The field of marketing is maturing. Substantially innovative products seem fewer and farther between. Innovation and success will come more and more from careful, well thought out new-product introductions.

References

Claycamp, H.J., and Liddy, L.E. 1969. "Prediction of New Product Performance: An Analytical Approach." *Journal of Marketing Research* 6(November):414-420.

Eskin, Gerald J. 1973. "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model." *Journal of Marketing Research* 10(May): 115-129.

Fourt, L.A., and Woodlock, J.W. 1960. "Early Prediction of Market Sales for New Grocery Products." *Journal of Marketing* 24 (October):31-38.

Little, John D.C. 1970. "Models and Managers: The Concept of a Decision Calculus." *Management Science* 16(April):B466-485.

Midgely, David F. 1977. *Innovation and New Product Marketing*. New York: Halsted Press.

Parfitt, J.H., and Collins, B.J.K. 1968. "Use of Consumer Panels for Brand-Share Prediction." *Journal of Marketing Research* 5(May): 131-145.

Urban, Glen L. 1970. "SPRINTER Mod III: A Model for the Analysis of New Frequently Purchased Consumer Products." *Operations Research* 18(September)-October):805-854.

Urban, Glen L., and Karash, Richard. 1971. "Evolutionary Model Building" *Journal of Marketing Research* 8(February):62-66.

Part V Early Sales-Based New-Product Forecasting Models

Once the product has been introduced to (regional or national) market, the early sales results can be utilized as the basis for a forecasting model. With limited sales data at hand, these models are concerned with estimating the growth of the product over time.

Part V contains two types of chapters for forecasting new-product sales with early sales data. The first set of chapters employs the diffusion-theory framework and focuses on the development of a product life-cycle curve. The underlying behavioral theory in the development of these models is that new-product acceptance is an imitation process. Most of these models have their roots and analogies in the models of epidemics, or biology and ecology, and serve the purpose of forecasting *product-class* sales for durable goods and novelty items. The best known first-purchase diffusion models of new-product acceptance in marketing are those of Bass (chapter 19) Fourt and Woodlock (1960), and Mansfield (1961). The seminal Bass model has been demonstrated successfully in retail service, industrial technology, agriculture, and consumer durable sectors (Bass, chapter 19 of this book; Nevers 1972; and Dodds 1973), while the Fourt and Woodlock model has been used to study success of certain grocery products (Fourt and Woodlock 1960). The Mansfield model and its revised forms, such as those proposed by Blackman (1974) and Fisher and Pry (1971), have been used in technological substitution studies of industrial innovations. The review chapter by Mahajan and Muller (chapter 18) assesses the state of the art of diffusion models of new-product acceptance, and the specific models are illustrated in two chapters, one by Bass (chapter 19) and the other by Midgley (chapter 20). The industry-utilization chapter by Lawrence and Lawton (chapter 22) demonstrates how diffusion models have been used successfully to model and forecast the product life cycles of a number of products.

The chapter by Rao (chapter 21) is concerned with predicting sales of new brands in a product category and employs brand-switching behavior as the basis for model development. Rao explains the Hendry System—Hendro-Dynamics—a technically less known commercial model to predict new-brand market shares. This model has been used in other stages of the new-product development process such as the pre-test-market stage. Further discussion of the utilization of brand-switching and stochastic choice models for forecasting new-product performance can be found, for exam-

ple, in Ehrenberg (1972), Bass (1974), Kalwani and Morrison (1977), Ehrenberg and Goodhart (1979), and Rubinson, Vanhonacker, and Bass (1980).

Additional Readings

Bass, Frank M. 1974. "The Theory of Stochastic Preference and Brand-Switching." *Journal of Marketing Research* 11 (February):1-20.

Blackman, A.W. 1974. "The Market Dynamics of Technological Substitutions." *Technological Forecasting and Social Change* 6 (February):41-63.

Dodds, W. 1973. "An Application of the Bass Model in Long Term New Product Forecasting." *Journal of Marketing Research* 10 (August): 308-311.

Dodson, J.A., and Muller E. 1978. "Models of New Product Diffusion through Advertising and Word-of-Mouth." *Management Science* 24 (November):1568-1578.

Ehrenberg, A.S.C. 1972. *Repeat Buying: Theory and Applications.* New York: American Elsevier Publishing Company, Inc.

Ehrenberg, A.S.C., and Goodhart, G.J. 1979. "The Switching Constant." *Management Science* 25 (July):703-705.

Fisher, J.C., and Pry, R.H. 1971. "A Simple Substitution Model of Technological Change." *Technological Forecasting and Social Change* 2 (May):75-88.

Fourt, L.A., and Woodlock, J.W. 1960. "Early Prediction of Market Success of New Grocery Products." *Journal of Marketing* 25 (October): 31-38.

Hendry, I. 1972. "The Three Parameter Approach to Long Range Forecasting." *Long Range Planning* 5 (March):40-45.

Kalwani, M.U., and Morrison, D.G. 1977. "A Parsimonious Description of the Hendry System." *Management Science* 23 (January):467-477.

Mansfield, E. 1961. "Technological Change and the Rate of Imitation." *Econometrica* 29 (October):741-766.

Nevers, J.V. 1972. "Extensions of a New Product Growth Model." *Sloan Management Review,* Winter, pp. 77-91.

Rubinson, J.R.; Vanhonacker, W.R.; and Bass, F.M. 1980. "On 'A Parsimonious Description of the Hendry System." *Management Science* 26 (February): 215-226.

18 Innovation Diffusion and New–Product Growth Models in Marketing

Vijay Mahajan and
Eitan Muller

The recent literature in new product planning has been concerned with developing better approaches for: (1) the generation and testing of new product concepts (Green and Wind 1975; Pekelman and Sen 1974; Shocker and Srinivasan 1974; Silk 1969; von Hippel 1978), (2) pretest market evaluation of new products (Claycamp and Liddy 1969; Silk and Urban 1978; Tauber 1977), (3) test-market evaluation of new products (Assmus 1975; Blattberg and Golanty 1978; Learner 1968; Urban 1970), (4) product-line strategy (Day 1977; Wind and Claycamp 1976), (5) product life cycle strategy (Dhalla and Yuspeh 1976), and diffusion models of new product acceptance (Mahajan and Peterson 1979b). This paper is concerned with diffusion models of new product acceptance.

The objective of a diffusion model is to represent the level of spread of an innovation among a given set of prospective adopters in terms of a simple mathematical function of time that has elapsed since the introduction of the innovation. The purpose of the model is to depict the successive increase in the number of adopters and predict the continued development of a diffusion process already in progress. In the product-innovation context, diffusion models are concerned with the spread of a new product from its manufacturers to ultimate users or adopters and focus on the development of a product life cycle curve (Kotler 1971; Wind 1974). The underlying behavioral theory in the development of these models is that new product acceptance is an imitation process (Rogers and Shoemaker 1971).

In recent years, a number of models have been developed to represent the spread of a new product in the marketplace. Most of these have their roots and analogies in the models of epidemics, or biology and ecology (Bailey 1957; Pearl 1925) and serve the purpose of forecasting sales for durable goods and novelty items. The objective of this paper is to review and assess the state-of-the-art of these modeling efforts. It is our hope that this review would provide the marketing managers and researchers with a

Reprinted from V. Mahajan and E. Muller, "Innovation Diffusion and New-Product Growth Models in Marketing," *Journal of Marketing* 43(October 1979):55–68, published by the American Marketing Association.

simple and systematic overview of the development of diffusion models of new product acceptance. In addition, we hope to provide answers to such questions as : What are the underlying assumptions of these models? Where have these models been applied? How are the proposed models different from each other? What are the shortcomings of these models? What directions need to be followed to make these models theoretically more sound and practically more effective and realistic? The marketing tradition of diffusion research has come on strong since the early 1960s (Rogers 1976). However, most of the diffusion models of new product acceptance have been developed in the last decade or so. A fair amount of the diffusion research in marketing has been and is being conducted in industry (e.g., Eastman Kodak, Mitre Corporation). Our presentation is restricted to the review and evaluation of the published literature.[1] Furthermore, the presentation is generally related to first-purchase diffusion models with some discussion of repeat purchase models.

The format of this exposition is as follows. First the general underlying structure employed in the development of diffusion models and the notation used are presented. Next, diffusion models are reviewed. Finally, issues related to the further development, validation, and evaluation of these models are discussed.

A General Diffusion Model Structure

The first-purchase diffusion models of new product acceptance assume that, in the product planning horizon being considered, there are no repeat buyers and purchase volume per buyer is one unit. In the development of these models, a model builder, in general, is concerned with the following question: At any time t in the diffusion process, if there are total $M(t)$ number of individuals in the market, and $\bar{N}(t)$ number of them can be considered as the potential customers (or the eventual number of buyers), how many of the potential customers will have bought the product by time t?

In other words, the analyst is concerned with (1) the total flow of customers and (2) the rate of flow of customers, across three distinct segments of the market. The first segment, the *untapped market,* consists of $(M(t) - \bar{N}(t))$ number of individuals. These are the individuals who are either uninformed about the product or due to a number of factors (such as high price of the product, general economic conditions, etc.) cannot be considered as the potential customers at time t. However, over time, individuals from this segment may transfer to the next segment, the *potential market,* consisting of $(\bar{N}(t) - N(t))$ number of potential customers at time t. The last segment, the *current market,* of course, includes the $N(t)$ number of adopters who have bought the product by time t. The summation of the

number of individuals in each segment totals the market, i.e.,

$$S_1(t) + S_2(t) + N(t) = M(t) \tag{18.1}$$

where

$$S_1(t) = M(t) - \bar{N}(t) \tag{18.2}$$

$$S_2(t) = \bar{N}(t) - N(t) \tag{18.3}$$

where $S_1(t)$ and $S_2(t)$ represent the number of customers in the first two segments, untapped market and potential market, respectively. Equations 18.1 to 18.3 characterize the customer flow in the diffusion process. This flow is summarized in figure 18-1. Furthermore, between time period t and $t + 1$, if $n(t)$ number of individuals transfer from the potential market segment to the current market segment, $\bar{n}(t)$ number of individuals transfer from the untapped market to the potential market and the total market increases by $m(t)$ (i.e., $m(t) = M(t + 1) - M(t)$), then at time $(t + 1)$ the number of individuals in the three market segments are given by the following equations:

$$S_1(t + 1) = S_1(t) + m(t) - \bar{n}(t) \tag{18.4}$$

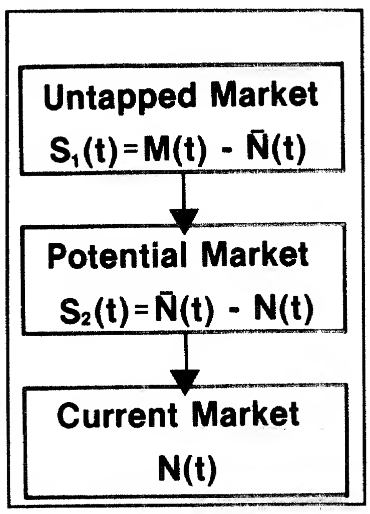


Figure 18-1. Customer Flow across the Three Market Segments in the Diffusion Process

$$S_2(t + 1) = S_2(t) + \bar{n}(t) - n(t) \tag{18.5}$$

$$N(t + 1) = N(t) + n(t) \tag{18.6}$$

Equation 18.4 implies that the number of individuals in the untapped market at time $(t + 1)$ is equal to number of individuals at time t plus the increase in the number of individuals in the total market minus the number of individuals who transfer from the untapped market to the potential market. Equation 18.5 suggests that the population of the potential market at time $(t + 1)$ is increased by the number of individuals who transfer from the untapped market to the potential market and decreased by the number of individuals who transfer from the potential market to the current market. Equation 18.6 reflects the increase in the number of adopters due to the $n(t)$ number of individuals who transfer from the potential market to the current market. If $s_1(t) = S_1(t + 1) - S_1(t)$ and $s_2(t) = S_2(t + 1) - S_2(t)$, equations 18.4 and 18.5 may be written as

$$s_1(t) = m(t) - \bar{n}(t) \tag{18.7}$$

$$s_2(t) = \bar{n}(t) - n(t) \tag{18.8}$$

Furthermore,

$$s_1(t) + s_2(t) + n(t) = m(t) \tag{18.9}$$

Equations 18.7 to 18.9 represent the rate of the flow of customers across the three segments of the diffusion process. In fact, the customer flow equations, 18.1 to 18.3, and the rate equations, 18.7 to 18.9, form the basis for all the proposed diffusion models. The development of these models basically involves the representation of the rate of customer flow across the segments in the diffusion process in terms of the transfer mechanisms. In general, five transfer mechanisms, which effect the flow of individuals from one segment to the next, may be considered. These are: (1) mass-media communication, (2) word-of-mouth communication, (3) other marketing efforts, (4) individual experience with the product, and (5) exogenous factors (e.g., general economic environment).

The available diffusion models differ in terms of the specific segments and the transfer mechanisms considered in the development. The following section provides a detailed comparison of these models.

Product Growth Models

The best-known, first-purchase diffusion models of new product acceptance in marketing are those of Bass (1969), Fourt and Woodlock (1960),

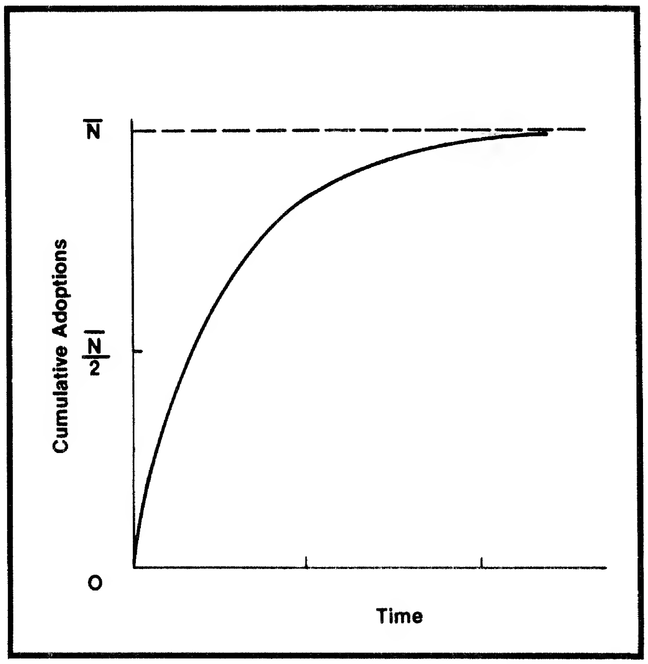


Figure 18-2. Modified Exponential-Function Diffusion Process

and Mansfield (1961). The Bass model has been successfully demonstrated in retail service, industrial technology, agriculture, and consumer durable sectors (Bass 1969; Dodds 1973; Nevers 1972), while the Fourt and Woodlock model has been used to study success of certain grocery products (Fourt and Woodlock 1960). The Mansfield model and its revised forms, such as those proposed by Blackman (1974) and Fisher and Pry (1971), have been used in technological substitution studies of industrial innovations.[2]

The Fourt and Woodlock model depicts the diffusion process, in terms of the number of customers who have bought the product by time t (i.e., $N(t)$), by a modified exponential curve (see figure 18-2). The Mansfield model, on the other hand, represents the diffusion process by a logistic curve (see figure 18-3). The Bass model synthesizes these two approaches and employs a generalized logistic curve which contains the Fourt and Woodlock, and Mansfield models as its special case. Related to the logistic curve is the Gompertz curve which also has been used to model the product growth (e.g., Hendry 1972) (see figure 18-4). The modified exponential curve, logistic curve, and Gompertz curve are the *basic* diffusion models of product growth.

In modeling the growth of the first-time buyers of a product, the basic diffusion models consider only two segments in the diffusion process—

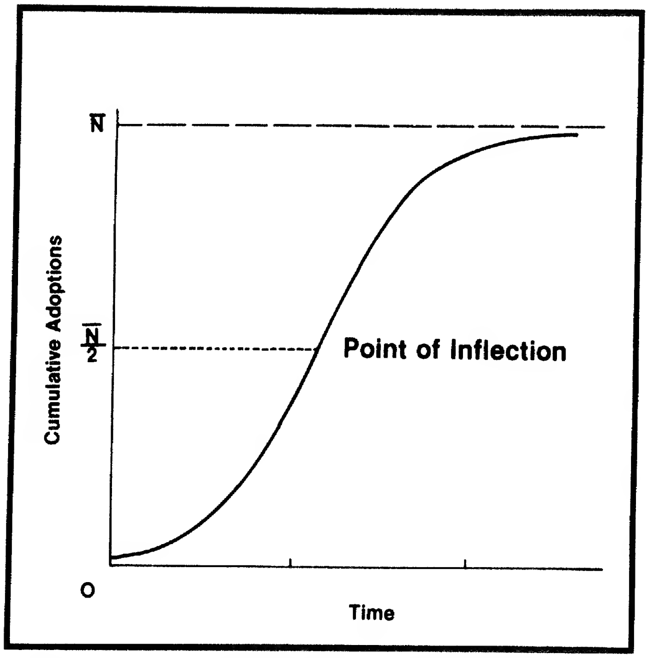


Figure 18–3. Logistic-Function Diffusion Process

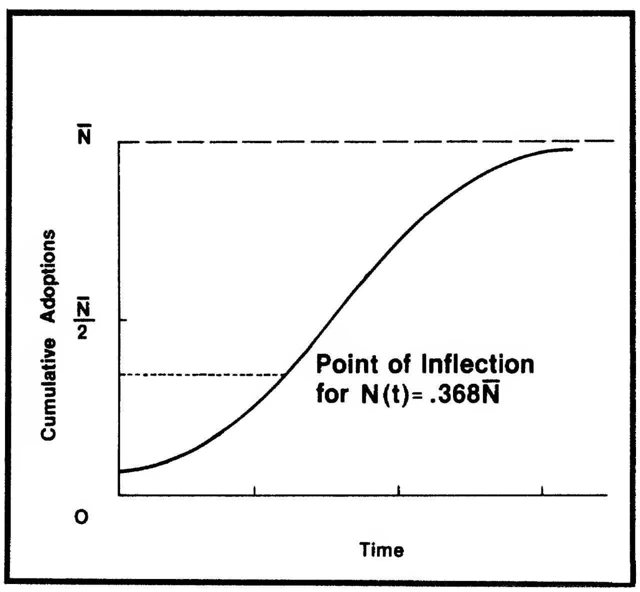


Figure 18–4. Gompertz's-Function Diffusion Process

potential market and current market—and two transfer mechanisms to influence the potential customers to adopt the product—mass-media communication and word-of-mouth communication. Furthermore, these models assume a *constant* total population of potential customers over the entire life of the product, i.e., $\bar{N}(t)$ is constant or $\bar{n}(t) = 0$. In other words, basic models are concerned with the following question: If there are $\bar{N}$ number of potential customers in the market, how many of the potential customers will have bought the product by time t? These models are *primarily concerned with modeling* $n(t)$, the flow of customers from the potential market to the current market.

A number of efforts have been made to extend these basic diffusion models. The extensions proposed by Bass (1978), Horsky and Simon (1978), Lilien and Rao (1978), Peterson and Mahajan (1978), and Robinson and Lakhani (1975) have been concerned with the inclusion of the effects of other transfer mechanisms in the basic diffusion models. However, these models also consider the customer flow across two segments of the diffusin process only—the potential market and current market—assume a constant total population of potential customers, i.e., $\bar{N}(t)$ is constant. Hence, these extensions also have been *primarily concerned with modeling* $n(t)$. On the other hand, extensions proposed by Chow (1967), Dodson and Muller (1978), Lackman (1978), and Mahajan and Peterson (1978) consider the customer flow across all three segments of the diffusion process. Furthermore, these extensions implicitly consider the following questions: If there are M number of individuals in the market and $\bar{N}(t)$ number of them can be considered as potential customers, how many of the potential customers will have bought the product by time (t)?

Note in the above question that these models assume that (1) M, number of individuals in the market, is constant, i.e., $m(t) = 0$; (2) $\bar{N}(t)$ total number of potential customers in the market, changes over time. In other words, these models are concerned with *modeling* $n(t)$, *the flow of customers from the potential market to the current market, as well as* $\bar{n}(t)$ (*or* $\bar{N}(t)$), *the flow of customers from the untapped market to the current market*. Finally, in extending the basic diffusion models to include the third segment, untapped market, Mahajan et al. (1979) relax the assumption that $M(t)$—the number of individuals in the market—is constant. In addition to modeling $n(t)$, $\bar{n}(t)$, they also are concerned with modeling the growth of the market, $m(t)$. The customer flow and the rate of customer flow equations for these models are given in appendix 18–A.[3]

Basic Product Growth Models

Considering the rate of customer flow between two segments of the diffusion process, the potential market and the current market, Bass (1969) proposed the following growth model for durables:

$$n(t) = p(\bar{N} - N(t)) + \frac{q}{\bar{N}}(\bar{N} - N(t))N(t) \qquad (18.10)$$

In the development of this model, Bass considered the underlying theory of innovative behavior which states that the diffusion of innovation is a communication process. In fact, he employed the so-called two-step flow of communication model (Robertson 1971). According to this model, the message about the product (through the mass media) is first picked up by a select few innovators who then "pass the word" to other members of the social system. Bass identified the constant p as the coefficient of innovation since the term containing the constant p in equation 18.10 represents adoptions without "interaction" with adopters. Similarly, he termed the constant q as the coefficient of imitation to reflect the word-of-mouth communication between adopters, $N(t)$, and potential adopters, $(\bar{N}\text{-}N(t))$. He further illustrated that for a successful product, the value of q should be greater than the value of p.

The interesting feature of the Bass model, equation 18.10, is that it unified the two earlier models proposed by Fourt and Woodlock (1960) and Mansfield (1961). Fourt and Woodlock demonstrated that the growth of certain grocery products can be modeled as the modified exponential curve, i.e.,

$$n(t) = p(\bar{N} - N(t)) \qquad (18.11)$$

Mansfield, on the other hand, found that the growth of certain technological innovations can be represented by an S-shaped or logistic curve, i.e.,

$$n(t) = bN(t)\,(\bar{N} - N(t)) \qquad (18.12)$$

Note $b = \frac{q}{\bar{N}}$ in the Bass model. Further note that the Fourt and Woodlock model, equation 18.11 assumes diffusion only through innovators and the Mansfield model, equation 18.12, assumes a pure imitation diffusion process. In a recent paper, Lekvall and Wahlbin (1973) further elaborate on the interpretation of these models. They question the validity of the two-step flow-of-communication model and suggest that in the communication process, each potential adopter is subjected to two different forms of influence—external and internal. External influence is the direct influence on the innovative behavior of an individual which, for example, the marketer of a new product exerts through various promotional activities such as mass media advertising, public exposure of the new product, or personal discussion with prospective adopters by professional salespeople. Internal influence, on the other hand, is the influence that the members of a social-system exert on one another as a result of social interaction. Furthermore,

these two influences may be operating simultaneously on an individiual at any time in the diffusion process. Using these two forms of the influence, Lekvall and Wahlbin suggest the following growth model:

$$n(t) = \frac{p}{\bar{N}}(\bar{N} - N(t)) + \frac{q}{\bar{N}}N(t)(\bar{N} - N(t)) \tag{18.13}$$

Note that the formulation suggested by equation 18.13 is exactly like the formulation suggested by equation 18.10, except that the constant p has been divided by $\bar{N}$. Unlike Bass, however, Lekvall and Wahlbin interpret the constant p as the coefficient of external influence representing the influence of promotional activities of a company. The constant q still represents the internal influence or the imitation effect. Commenting further on these two coefficients, Lekvall and Wahlbin write:

> In most real world situations external and internal influence will interact to produce a certain outcome of a diffusion process. However, the relative strength of these two factors are not likely to be the same in all situations. In some cases the dominating influential power may emanate from sources outside the set of (prospective) adopters, whereas in other situations the influence that the members of this set exert on each other may be more important. Furthermore, theoretical considerations as well as empirical evidence indicate that the nature of the innovation is a crucial factor underlying such differences. For example, whereas a new kind of fertilizer may be subject to much debate among farmers in a community—and thus show a relatively high intensity of internal influence—this is not likely to be the case for a new kind of orange juice or a new brand of household soap. Likewise, whereas the buyer of a new car model will probably initiate substantial internal influence just by using the car, the adoption of disposable toothbrushes, for example, is not likely to yield any considerable demonstration effect.

Through a simulation process Lekvall and Wahlbin generate different shapes of diffusion curves reflecting different intensity levels of p and q.

Following Lekvall and Wahlbin's interpretation of the constants p and q, it is clear that the modified exponential curve or the Fourt and Woodlock model, equation 18.11, represents the growth of a product which is mainly caused by the external influence or the promotional activities of a company. The logistic curve, equation 18.12 on the other hand, represents the growth of a product which is mainly caused by the internal influence.

Related to the logistic curve is the Gompertz curve which has also been used to model the diffusion process:

$$n(t) = bN(t)(\ln\bar{N} - \ln N(t)) \tag{18.14}$$

As evident from equation 18.14, like the logistic curve, the Gompertz

curve considers only internal influence. However, unlike the logistic curve, the Gompertz curve expresses the difference between the total number of individuals in the potential market and current number of adopters in terms of their natural logarithms. A further observation on the difference between these two growth curves is in terms of the stage at which the maximum rate of growth, $n(t)$, is reached. Setting the derivative of equation 18.12 with respect to $N(t)$ equal to zero, one finds that $N(t) = .5\bar{N}$. Setting the derivative of equation 18.14 equal to zero, one finds that $N(t) = e^{-1}\bar{N} = .37\bar{N}$. Hence, the maximum rate of growth is reached when the total number of adopters is about 37% of its potential market by the Gompertz curve, and when the total number is 50% of the potential market by the logistic curve. To put it in other words, if the Gompertz curve is used the number of buyers at any time t, $n(t)$, cannot be greater than 37% of the potential market. Furthermore, the logistic curve is symmetrical, i.e., it gives the same value of $n(t)$ *for* $N(t) = .5\bar{N} + k$ or $N(t) = .5\bar{N} - k$, where k is constant. On the other hand, the Gompertz curve attains its maximum rate of growth at an earlier stage and maintains a more nearly constant rate of growth later on, than does the logistic curve. The application of the Gompertz curve has been demonstrated by Hendry (1972) to model the growth of certain durables in the United Kingdom.

Extensions of the Basic Models of Product Growth

The major criticism of the basic diffusion models of new product acceptance is that they are of little use to the new product manager since they consider diffusion as a function of time only and the marketing program of a company does not enter explicitly as a variable inhibiting the evaluation of the effect of different marketing strategies on the product growth. Since all these growth curves essentially contain three parameters, i.e., coefficient of external influence, coefficient of internal influence, and the total number of potential customers, marketing efforts may be incorporated into these models by representing these parameters as a function of relevant variables, i.e.,

$$a(t) = A(S(t)) \tag{18.15}$$

$$b(t) = B(S(t)) \tag{18.16}$$

$$\bar{N}(t) = N(S(t)) \tag{18.17}$$

where $a(t)$ is the coefficient of external influence at time t, $b(t)$ is the coefficient of internal influence at time t, $\bar{N}(t)$ is the total number of potential

customers at time t, and $S(t)$ is a vector of all relevant marketing decision variables. By definition, then, the basic diffusion models of new product growth, i.e., modified exponential curve, logistic curve, generalized logistic, and Gompertz curve, assume that the marketing program of a company remains constant over the entire life of the product growth.[4] The various extensions proposed, however, do not offer a unified theoretical framework to include the effects of marketing variables on the product growth. For this reason, each individual extension will be presented separately. First, consider the extensions to the generalized logistic curve, equation 18.10 or 18.13.

(a) Robinson and Lakhani (1975) point out that since for consumer durables, the coefficient of external influence (they call it coefficient of innovation as suggested by Bass 1969) has been found to be very small it is the coefficient of internal influence which should be developed as a function of marketing decision variables. They argue that innovators are only a dominant factor in the marketplace during the short period required to achieve the first several percent of market penetration. Therefore, if the coefficient of internal influence is developed as a function of marketing decision variables such as advertising and price, the diffusion model will enable management to evaluate the effect of a certain marketing program on the growth of the product. In their work, Robinson and Lakhani specifically study the effect of price on the product and assume the following values of the three parameters:

$$\begin{aligned} a &= \text{constant} \\ b(t) &= b_1 \exp(-EP(t)) \\ \bar{N} &= \text{constant} \end{aligned} \tag{18.18}$$

where b_1 is a constant, E is price elasticity, and $P(t)$ is price at time t.

In a recent paper, incorporating the effect of price in the diffusion models, Bass (1978) argues that as a result of learning, costs and prices for new durables decline over time and consequently affect the growth of the product. It is shown by Mahajan and Peterson (1979b) that Bass assumes the following values of the parameters.

$$\begin{aligned} a(t) &= p \cdot k \cdot (N(t))^{\lambda\alpha} \\ b(t) &= \frac{q}{\bar{N}} \cdot k \cdot (N\text{t}))^{\lambda\alpha} \\ \bar{N} &= \text{constant} \end{aligned} \tag{18.19}$$

where p and q are coefficients of innovation and imitation, respectively (see

equation 18.10), k is a constant, α is elasticity of demand, and λ is the learning parameter in the experience curve of production quantity and price. Since price is assumed to be inversely related to the quantity sold, equations 18.19 can also be expressed as a function of price.

Unlike the Robinson and Lakhani model, equation 18.18, the new Bass model suggests that both the coefficients of external influence and internal influence change over time and become constant only when $N(t) = \bar{N}$. This development is consistent with the observation made by Kotler (1971) that as more units are sold the coefficient a (and possibly b) should change over time. Bass has tested his new model for six durables.[5]

(b) Incorporating the effect of advertising into the diffusion model, Horsky and Simon (1978) believe that in the context of new products the primary effect of advertising is to be a direct tool for disseminating information about the existence of the new product. Therefore, expression of the coefficient of internal influence as a function of advertising will result in an effect of second order. They suggest that a correct specification of a diffusion model should express the coefficient of external influence as a function of advertising expenditures. In their model application to a new telephonic banking system, Horsky and Simon assume the following values of the parameters:

$$\begin{aligned} a(t) &= a_1 + a_2 A(t) \\ b &= \text{constant} \\ \bar{N} &= \text{constant} \end{aligned} \tag{18.20}$$

where a_1 and a_2 are constants and $A(t)$ is the level of advertising expenditures.

(c) Modeling the spread of two ethical drugs aimed at a certain class of doctors, Lilien and Rao (1978) consider the effect of personal selling on the product growth. They point out that one of the most important components of the marketing mix employed by pharmaceutical companies is "detailing"—i.e., personal selling by a force of "detailmen" who visit doctors and describe their portfolio of products, provide free samples, literature, etc. For a new product, the impact of detailing is augmented by the word-of-mouth effect that occurs when doctors first prescribing the product find it satisfactory and recommend it to their colleagues. In their model application, Lilien and Rao consider the following values of the parameters in the first-purchase diffusion model:

$$a(t) = a_1 d(t) + a_2 d^2(t) \tag{18.21}$$

$$\begin{aligned} b &= \text{constant} \\ \bar{N} &= \text{constant} \end{aligned}$$

where a_1 and a_2 are constants and $d(t)$ is the level of detailing at time t.

(d) Studying the simultanous growth of related products, Peterson and Mahajan (1978) argue that new products are not introduced into a vacuum nor do they exist in isolation. Other products exist in the market, and these may have an influence—positive or negative—upon the growth of a new product. They specifically consider four product relationships—independent, complementary, contingent, and substitute. They implicitly incorporate the effect of product relationship into the coefficient of internal influence. For example, in the case of two *complementary* products, the parameters for the *first* product are:

$$
\begin{aligned}
a &= \text{constant} \\
b(t) &= b_1 + c_1 \frac{N_1(t)}{N_2(t)} \\
\bar{N} &= \text{constant}
\end{aligned}
\tag{18.22}
$$

where b_1 represents the coefficient of internal influence between the potential adopters and adopters of product 1, and c_1 represents the positive influences of the adopters of products 2 on the potential adopters of product 1. $N_2(t)$ and $N_1(t)$ are total number of adopters of products 2 and 1, respectively.

(e) Considering the effect of marketing efforts and exogeneous variables on the product growth, Mahajan and Peterson (1978) suggest that number of potential customers, i.e., $\bar{N}(t)$. In fact, they suggest that $\bar{N}(t)$ changes over time and may even be affected by the exogeneous variables such as general economic conditions, changing characteristics of the individuals in the market, technological changes, government actions, etc. The objective of advertising and distribution channels is to increase the awareness and availability of a new product. The effect of this activity should result in an expansion of the total potential market. Their arguments also are supported by Horsky and Simon (1978) who, justifying the development of their model (extension (b) above), suggest that due to increased competition and reduction in production costs as a result of learning, price reduction will place the product within the budgetary limitations of a great number of potential customers, thus expanding the total potential market of the product. Furthermore, they argue that the objective of product modifications and technological changes is to tailor the product to different segments as the interests of those segments become apparent. The effect of this activity is also likely to manifest itself in an expansion of the potential market, $\bar{N}(t)$.

Figure 18-5 illustrates the dynamic growth model suggested by Mahajan and Peterson (1978, 1979). The following features of the dynamic model can be noted in this figure:

Market potential of the new product is increasing, and the market potential curve is distinctly different, at least in the early stage of the product growth, from the product growth curve.

The difference between the total number of potential adopters, given by the market potential curve, and the actual number of adopters, given by the product growth curve, at any time t decreases with time, and ultimately,

The market potential curve coincides with the product growth curve.

In their dynamic growth model development, Mahajan and Peterson (1978) consider the following values of the parameters:

$$a = \text{constant}$$

$$b = \text{constant}$$

$$\bar{N}(t) = f(\text{marketing efforts, exogenous factors}) \qquad (18.23)$$

Recalling equations 18.7 to 18.9, note here that Mahajan and Peterson consider the flow of customers across all the three segments of the diffusion

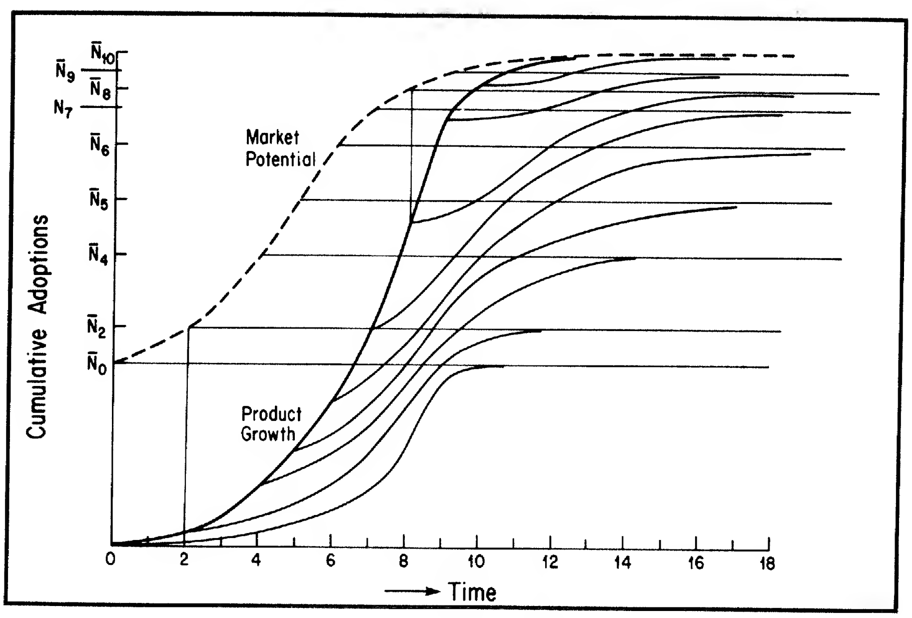


Figure 18-5. Dynamic Product Growth Patterns

process. The use of their model involves modeling $n(t)$ as well as $\bar{n}(t)$. In the empirical illustrations of their model for consumer durables, they consider the effect of the construction activities, defined in terms of the number of housing starts, on the product growth (Mahajan et al. 1979). That is, considering that the total market, $M(t)$, changes over time, and recalling equations 18.7 to 18.9, they specify the customer flow across the three market segments of the diffusion process by the following equations:

$$m(t) = (h_1 + h_2M(t))(\bar{M} - M(t)) \tag{18.24a}$$

$$n(t) = k_2m(t) \tag{18.24b}$$

$$n(t) = bN(t)(\bar{N}(t) - N(t)) \tag{18.24c}$$

where equation 18.24*a* represents the growth of housing starts in terms of the generalized logistic curve (h_1, h_2, and $\bar{M}$ are constants), equation 18.24*b* considers the increase in the number of potential customers (k_2 is a constant), and equation 18.24*c* represents the growth of durables by the logistic curve. From equations 18.24 it is obvious that they use the following values of the diffusion parameters:

$$\begin{aligned} a &= 0 \\ b &= \text{constant} \\ \bar{N}(t) &= k_1 + k_2M(t) \end{aligned} \tag{18.25}$$

(f) Incorporating the effect of advertising into the growth models and considering the customer flow across all the three segments of the diffusion process, Dodson and Muller (1978) suggest that since $\bar{N}(t)$, eventual number of buyers at time t, represents the number of potential buyers who are aware of the product, it is $\bar{N}(t)$ which should be expressed as a function of the level of advertising expenditure. In fact, they suggest two mechanisms by means of which unaware individuals may become aware of the product: word-of-mouth communication between product aware and unaware individuals, and advertising. As shown in appendix 18B, the Dodson and Muller model is completely specified by the following rate equations:

$$\begin{aligned} \bar{n}(t) &= k_1\bar{N}(t)(M - \bar{N}(t)) + k_2(M - \bar{N}(t)) \\ n(t) &= a(\bar{N}(t) - N(t)) \end{aligned} \tag{18.26}$$

Note in equations 18.26 that Dodson and Muller use the modified exponential curve to model the growth of the product and the logistic curve to model

the increase in the total number of potential customers. The constants k_1 and k_2 in equations 18.26 are interpreted as follows. Since the total number of customers in the potential market, $\bar{N}(t)$, consists of individuals who are aware of the product, $(M - \bar{N}(t))$ represents the individuals who are unaware of the product. Hence the term containing k_2 represents the increase in the number of informed individuals or the potential customers due to external influence such as *advertising*. Similarly, the term containing k_1 represents the increase in the number of informed individuals due to word-of-mouth communication between the informed, $\bar{N}(t)$, and uninformed individuals, $(M - \bar{N}(t))$. Hence, k_1 and k_2 are coefficients of internal influence and external influence, respectively, for the potential market growth. To summarize, Dodson and Muller assume the following values of the diffusion parameters:

$$a = \text{constant}$$

$$b = 0 \tag{18.27}$$

$$\bar{N}(t) = f(\text{advertising, word-of-mouth})$$

Note here the similarities between the dynamic growth model, equations 18.23, suggested by Mahajan and Peterson and the model suggested by Dodson and Muller, equations 18.27. The model suggested by Mahajan and Peterson is a general dynamic growth model and the model suggested by Dodson and Muller presents a mechanism to incorporate the advertising variable into the dynamic growth model.

The above are the seven reported extensions of the generalized logistic curve.[6] Consider, now, the two identified extensions to the Gompertz curve, equation 18.14. Both of these are related to the inclusion of the effect of marketing efforts and exogeneous factors on the total number of potential buyers, $\bar{N}(t)$. In other words, these extensions consider the flow of customers across all the three segments. However, they assume a constant population of individuals in the market. A brief description of these extensions follows:

(a) Studying the growth of a new plastic product in the automotive industry, Lackman (1978) uses the following parameters:

$$b = \text{constant} \tag{18.28}$$

$$\bar{N}(t) = \bar{N} \cdot (\frac{ZB(t)}{SC(t)})^k \tag{18.28}$$

where k is a constant and $ZB(t)/SC(t)$ is the ratio of corporate profits to corporate sales in the automotive industry. The profit-sales variable is in-

cluded to reflect the fact that product users shift to new product quickly when profitability is high.

(b) Examining the natural growth of computers Chow (1967) argues that the number of computer adoptions have been influenced by the technological changes and price reductions. He implicitly assumes the following formulation to represent this effect:

$$\bar{N}(t) = B_0(P(t))^{-B_1} \tag{18.29}$$

where B_0 and B_1 are constants and $P(t)$ is the price.

The summary of these extensions is provided in table 18–1.

Issues

After a review of underlying assumptions, applications, differences, and shortcomings of various diffusion models of new product growth, it is important to indicate certain key issues which must be addressed in order to make these models theoretically more sound and practically more effective and realistic. These issues relate to the purpose, theory, model formulation, and empirical validation and evaluation of these models.

Purpose

The diffusion models have been developed to estimate new product sales. However, in order to determine the exact shape of the diffusion curve, some early data on the actual diffusion process of the product are required. By the time a new product is launched, investment has been committed and, to that extent, forecasts become less useful. Therefore, for the diffusion models to be useful in developing prelaunch forecasts, various procedures based on laboratory tests and market research need to be examined to estimate model parameters. These approaches may involve the use of parameters derived from an empirical analysis of past new product introductions and experience surveys. The former approach has been illustrated by Dodds (1973) and Lilien and Rao (1978). Using the basic Bass model, equation 18.10, to study the growth of cable television, Dodds (1973) successfully employed the total market potential estimate, i.e., $\bar{N}$, developed for color television. Similarly, Lilien and Rao (1978) employed the estimates derived from a fully diffused ethical drug to study the spread of a new ethical drug. These approaches seem very promising. The authors understand that Eastman Kodak is working with a number of products to estimate general values for the coefficient of external influence and the coefficient of internal influence in the basic diffusion models.

Included in the parameters estimation issue of these models is the fact

Table 18–1
First-Purchase Diffusion Models of New-Product Acceptance

Work by:	*Coefficient of Internal Affluence*	*Coefficient of External Influence*	*Total Number of Potential Customers*
Basic Models			
Bass (1969)	Constant	Constant	Constant
Fourt and Woodlock (1960)	0	Constant	Constant
Mansfield (1961)	Constant	0	Constant
Gompertz curve (for example, Hendry 1972)	Constant	0	Constant
Lekvall and Wahlbin (1973)	Constant	Constant	Constant
Extensions			
Robinson and Lakhani (1975)	*f*(price)	Constant	Constant
Horsky and Simon (1978)	Constant	*f*(advertising)	Constant
Lilien and Rao (1978)	Constant	*f*(personal selling)	Constant
Bass (1978)	*f*(demand elasticity, learning parameters, price)	*f*(demand elasticity, learning parameters, price)	Constant
Peterson and Mahajan (1978)	*f*(product relationships)	Constant	Constant
Mahajan and Peterson (1978)	Constant	Constant	*f*(all relevant variables)
Mahajan et. al. (1979)	Constant	0	*f*(housing starts)
Dodson and Muller (1978)	0	Constant	*f*(advertising, word-of-mouth)
Chow (1967)	Constant	0	*f*(price)
Lackman (1978)	Constant	0	*f*(profit/sales)

that one is working with very few data points and multicollinearity among the variables. One may obtain parameter estimates which are unstable with high standard errors. This may require consideration of approaches such as ridge regression to handle multicollinearity (Mahajan, Jain, and Bergier 1977) or the feedback/adaptive approaches which automatically adjust and update the initial parameter estimates to changing data patterns (Mahajan, Bretschneider, and Bradford 1980). Furthermore, the use of feedback/adaptive approaches may also assist in dealing with the range of the values of the decision variables not encountered in the earlier data.

Theory

From the review of the extensions to the basic diffusion models, it is apparent that each individual effort has tried to incorporate only one relevant variable into the model; and different arguments have been made to incor-

porate the same variable either in the coefficient of internal influence, coefficient of external influence, or the total number of potential customers. For example, consider the variable advertising. Horsky and Simon (1978) believe that it should be included in the coefficient of external influence. Dodson and Muller (1978) suggest that its effect should be included in the total market potential. Similar is the case with price. Robinson and Lakhani (1975) consider its effect on the coefficient of internal influence. Mahajan and Peterson (1978), Horsky and Simon (1978), and Chow (1967) have argued that its effect should be considered in the total market potential. In spite of all these extensions, a unified theory to incorporate the marketing variables (as well as exogenous factors) into diffusion models is not apparent. Furthermore, these extensions have dealt only with price and promotion. Some work is underway in geography by Brown (1979) to examine the effect of the distribution network on the diffusion process. There remains a woeful lack of accomplishment in this domain.

Model Formulation

Although some significant contributions have been made to the basic diffusion models, a number of further refinements still remain. First is the inclusion of repeat or replacement sales. Of all the models summarized in table 18-1,[7] only the models proposed by Dodson and Muller (1978) and Lilien and Rao (1978) have considered this issue.[8] However, both the models are restricted to the first-repeat sales. Furthermore, none of the models has included impulse purchase.

The most important refinement, however, is related to the underlying assumption of a complete social network. These models assume that all of the adopters, $\bar{N}(t)$, (or informed individuals in case of the Dodson and Muller model), interact with all the nonadopters ($\bar{N}(t) - N(t)$), or (uninformed individuals). This assumption is unrealistic since not all of the adopters or informed individuals are communicators of the product. Furthermore, these models have implicitly combined the effect of two distinct transfer mechanisms--individual experience and word-of-mouth—in the sense they assume that individual experience with the product is always communicated positively through word-of-mouth. It is very likely that the communicators of the product experience may transfer favorable, unfavorable, or indifferent messages through word-of-mouth. In other words, the fifth transfer mechanism—individual experience with the product—may not only influence the adopter's decision to repurchase, but may also influence potential adopters positively or negatively. What is now required is a comprehensive diffusion model which considers the customer flow across the untapped, potential, and current-market (consisting of triers or first-purchasers and

second, third . . . nth repeat purchasers) identifying the communicators of the product and distinguishing the type of message conveyed. A start in this direction has been made in the innovative behavior model proposed by Midgley (1977).

Finally, diffusion models should possess parameters which are not abstract, but show the exact relationship between the product growth and marketing decision variable so that the effect of different marketing strategies on the product growth can be assessed. Furthermore, most of the current diffusion models are deterministic in that they provide point estimates. Hence, further extensions of these models should examine stochastic model formulations to provide a probabilistic range of sales estimates.

Empirical Validation and Evaluation

Regardless of the fact that some of the models have been available for some time, there has been relatively little material on their validity and reliability. There is a distinct "need" to know when the models "work" and when they do not. Furthermore, in order to establish the superiority of these models they need to be compared with other forecasting techniques to establish their (1) short- and long-run forecasting accuracy, (2) ability to identify future turning points, (3) time and cost requirements for model development and implementation, and (4) diagnostic power (Chambers, Mullick, and Smith 1971; Larreche and Montgomery 1977; Makridakis and Wheelright 1977; Wind 1974).

Conclusions

The objective of this paper has been to review and assess the state-of-the-art of diffusion models of new product acceptance. After reviewing the first-purchase models, a number of issues related to their further development and validation were presented. There have been some significant contributions to the development of these models but further development and continuous validation are needed to establish them as the useful tools of new product planning.

Notes

1. There are at least 18 different disciplines which study the diffusion of innovation, looking at its different aspects and different contexts. The review in this paper is restricted to marketing literature with some key bor-

rowings from economics literature. Although Rogers and Shoemaker (1971) and Robertson (1971) provide excellent reviews of the state-of-the-art until 1971, some recent reviews, for example, by Brown (1979) in geography, Hurter and Rubenstein (1978) in technological substitution, and Rogers (1976) in communication, are very informative. The spatial diffusion models of new product acceptance also are not reviewed. Brown (1979) and Mahajan and Peterson (1979*a*) provide a review of these modeling efforts.

2. The technological forecasting literature identifies the diffusion process mainly as a substitution process by which the adoption of a new product or technology spreads and grows to replace an existing product or technology. Linstone and Sahal (1976) provide an excellent review of this literature.

3. Some further possible extensions of the basic models are discussed in Bernhardt and MacKenzie (1972). In the communication literature, Rogers and Shoemaker (1971) have suggested that the noncumulative adoption curve is a normal distribution. However, the cumulative normal distribution also yields a logistic curve, one of three basic models presented here.

4. It should be noted here that since the basic diffusion models focus on the development of a product life cycle curve, incorporation of marketing variables into the three parameters of the diffusion model complements the arguments advanced by Dhalla and Yuspeh (1976) that the product life cycle is not given, but can be controlled by the marketing decision variables.

5. In the development of his model, Bass derives

$$\bar{N}(t) = (K(1 - \lambda\alpha))^{\left(\frac{1}{1-\lambda\alpha}\right)}$$

where $K = c\left(\dfrac{\alpha}{\alpha - 1}\, c_1\right)^{-\alpha}$ and c, c_1 are constants. Since λ and α are assumed to be constant, $\bar{N}(t)$ is also constant. Hence, in the calibration and the application of his model, Bass uses a constant total market potential, i.e. constant value of $\bar{N}(t)$.

6. Hernes (1976) also has proposed an extension to the generalized logistic curve. He argues that because of the changing characteristics of the population of the social system, and the innovation (technological changes), the coefficients of external and internal influence should change over time. However, he does not explicitly include the effect of exogenous variables on these diffusion parameters. Studying the spread of TV ownership in Norway, Hernes used the following values of the parameters.

$$a(t) = a_1 \cdot a_2^t$$

$$b(t) = b_1 \cdot b_2^t$$

$$\bar{N} = \text{constant}$$

where a_1, a_2, b_1, and b_2 are constants. Hernes believes that his model is very flexible in the sense that different types of diffusion curves can be obtained by considering different values of the four constants. For example, if a_2, = b_2 = 1, the generalized logistic formulation is obtained.

7. Other repeat purchase models include those suggested by Fourt and Woodlock (1960), Eskin (1973), and Parfitt and Collins (1968). As compared to Fourt and Woodlock and Eskin, the Parfitt and Collins model is designed for long-run share predictions for a new brand introduced into an established market, rather than for a new product innovation. This model has been further refined by Shoemaker and Staelin (1976). Although none of the three models really use any of the innovation diffusion concepts, they do suggest the type of diffusion curves which may be employed to study the repeat purchases.

8. Dodson and Muller illustrated that the advertising models proposed by Glaister (1974), Gould (1970), Nerlove and Arrow (1962), and Nicosia (1966) can be shown to be special cases or related to their model.

References

Assmus, G. (1975), "NEWPROD: The Design and Implementation of a New Product Model," *Journal of Marketing Research,* 39 (January), 16–23.

Bailey, N.T.J. (1957), *Mathematical Theory of Epidemics,* New York: Hafner Publishing Company.

Bass, F.M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science,* 15 (January), 215–227.

——— (1978), "An Integration of the New Product Growth Model with the Experience Cost Function and Optimal Pricing," presented at the TIMS/ORSA National Meeting in New York City, May.

Bernhardt, L., and K.E. MacKenzie (1972), "Some Problems in Using Diffusion Models for New Products," *Management Science,* 18 (October), 187–200.

Blackman, A.W. (1974), "The Market Dynamics of Technological Substitutions," *Technological Forecasting and Social Change,* 6 (February), 41–63.

Blattberg, R., and J. Golanty (1978), "TRACER: An Early Test Market Forecasting and Diagnostic Model for New Product Planning," *Journal of Marketing Research,* 15 (May), 192–202.

Brown, L.A. (1979), "Innovation Diffusion: A New Perspective,' Studies in Diffusion of Innovation, Discussion Paper No. 60, Department of Geography, The Ohio State University.

Chambers, J.C., S.K. Mullick, and D.D. Smith (1971), "How to Choose the Right Technique," *Harvard Business Review,* 49 (July–August), 45–74.

Chow, G.C. (1967), "Technological Change and the Demand for Computers," *American Economic Review,* 57 (December), 1117–1130.

Claycamp, H.J., and L.E. Liddy (1969), "Prediction of New Product Performance: An Analytical Approach," *Journal of Marketing Research,* 6 (November), 414–420.

Day, G.S. (1977), "Diagnosing the Product Portfolio," *Journal of Marketing,* 41 (April), 29–38.

Dhalla, N.K., and S. Yuspeh (1976), "Forget the Product Life Cycle Concept," *Harvard Business Review,* 54 (January), 102–112.

Dodds, W. (1973), "An Application of the Bass Model in Long-Term New Product Forecasting," *Journal of Marketing Research,* 10 (August), 308–311.

Dodson, J.A., and E. Muller (1978), "Models of New Product Diffusion through Advertising and Word-of-Mouth," *Management Science,* 15 (November), 1568–1578.

Eskin, G.J. (1973), "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model," *Journal of Marketing Research,* 10 (May), 115–129.

Fisher, J.C., and R.H. Pry (1971), "A Simple Substitution Model for Technological Change," *Technological Forecasting and Social Change,* 2 (May), 75–88.

Fourt, L.A., and J.W. Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* 25 (October), 31–38.

Glaister, S. (1974), "Advertising Policy and Returns to Scale," *Economica,* 41 (May), 139–156.

Gould, J.P. (1970), "Diffusion Processes and Optimal Advertising Policy," in *Microeconomic Foundation of Employment and Inflation Theory,* E.S. Phelps, et al., eds., New York: W.W. Norton.

Green, P.E., and Y. Wind (1975), "New Ways to Measure Consumer Judgments," *Harvard Business Review,* 53 (July–August), 107–117.

Hendry, I. (1972), "The Three Parameter Approach to Long Range Forecasting," *Long Range Planning,* 5 (March), 40–45.

Hernes, G. (1976), "Diffusion and Growth—The Non-Homogeneous Case," *Scandinavian Journal of Economics,* 78, 427–436.

Horsky, D., and L.S. Simon (1978), "Advertising in a Model of New Product Diffusion," presented at the TIMS/ORSA National Meeting in New York City, May.

Hurter, A.P., and A.H. Rubenstein (1978), "Market Penetration by New Innovations: The Technological Literature," *Technological Forecasting and Social Change,* 11, 197–221.

Kotler, P. (1971), *Marketing Decision Making: A Model Building Approach,* New York: Holt, Rinehart and Winston.

Lackman, C.L. (1978), "Gompertz Curve Forecasting: A New Product Application," *Journal of the Market Research Society,* 20 (January), 45-47.

Larreche, J., and D.B. Montgomery (1977), "A Framework for the Comparison of Marketing Models: A Delphi Method," *Journal of Marketing Research,* 14 (November), 487-498.

Learner, D.B. (1968), "Profit Maximization through New-Product Marketing Planning and Control," in *Applications of the Sciences in Marketing Management,* F.M. Bass, C.W. King, and E.A. Pessemier, eds., New York: John Wiley and Sons, Inc.

Lekvall, P., and C. Wahlbin (1973), "A Study of Some Assumptions Underlying Innovation Diffusion Functions," *Swedish Journal of Economics,* 75, 362-377.

Lilien, G.L., and A.G. Rao (1978), "A Marketing Promotion Model with Word of Mouth Effect," Working Paper 976-78, Cambridge, Mass.: Sloan School of Management, Massachusetts Institute of Technology.

Linstone, H.A., and D. Sahal (197), *Technological Substitution,* New York: American Elsevier Publishing Company.

Mahajan, V., and R.A. Peterson (1978), "Innovation Diffusion in a Dynamic Potential Adopter Population," *Management Science,* 15 (November), 1589-1597.

——— (1979*a*), "Integrating Time and Space in Technological Substitution Models," *Technological Forecasting and Social Change* (forthcoming).

——— (1979*b*), "First Purchase Diffusion Models of New Product Acceptance," *Technological Forecasting and Social Change* (forthcoming).

Mahajan, V., A.K. Jain, and M. Bergier (1977), "Parameter Estimation in Marketing Models: An Application of Ridge Regression," *Journal of Marketing Research,* 14 (November), 586-591.

Mahajan, V., S.I. Bretschneider, and J.W. Bradford (1980), "The Feedback Approaches to Modeling Structural Shifts in Market Response," *Journal of Marketing* (forthcoming).

Mahajan, V., R.A. Peterson, A.K. Jain, and N. Malhotra (1979), "A New Product Growth Model with a Dynamic Market Potential," *Long Range Planning* (forthcoming).

Makridakis, S., and S.C. Wheelright (1977), "Forecasting: Issues and Challenges for Marketing Management," *Journal of Marketing,* 41 (October), 24-38.

Mansfield, E. (1961), "Technical Change and the Rate of Imitation," *Econometrica,* 29 (October), 741-766.

Midgley, D.F. (1977), *Innovation and New Product Marketing,* New York: Halsted Press.

Nerlove, M., and K.J. Arrow (1962), "Optimal Advertising Policy under Dynamic Conditions," *Economica,* 29 (May), 129–142.

Nevers, J.V. (1972), "Extensions of a New Product Growth Model," *Sloan Management Review,* 13 (Winter), 78–79.

Nicosia, F.M. (1966), *Consumer Decision Processes,* Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Parfitt, J.H., and B.J.K. Collins (1968), "Use of Consumer Panels for Brand-Share Predictions," *Journal of Marketing Research,* 5 (May), 131–145.

Pearl, R. (1925), *Studies in Human Biology,* Baltimore: Williams and Wilkins.

Pekelman, D., and S.K. Sen (1974), "Mathematical Programming Models for the Determination of Attribute Weights," *Management Science,* 20 (April), 1217–1229.

Peterson, R.A., and V. Mahajan (1978), "Multi-Product Growth Models," in *Research in Marketing,* J. Sheth, ed., Greenwich, Conn.: JAI Press.

Robertson, T.S. (1971), *Innovative Behavior and Communication,* New York: Holt, Rinehart and Winston.

Robinson, V., and C. Lakhani (1975), "Dynamic Price Models for New Product Planning," *Management Science,* 21 (June), 1113–1132.

Rogers, E.M. (1976), "New Product Adoption and Diffusion," *Journal of Consumer Research,* 2 (March), 290–301.

Rogers, E.M., and F.F. Shoemaker (1971), *Communication of Innovations: A Cross-Cultural Approach,* New York: The Free Press.

Shocker, A.D., and V. Srinivasan (1974), "A Consumer-Based Methodology for the Identification of New Product Ideas," *Management Science,* 20 (February), 921–937.

Shoemaker, R., and R. Staelin (1976), "The Effects of Sampling Variation on Sales Forecasts for New Consumer Products," *Journal of Marketing Research,* 13 (May), 138–143.

Silk, A.J. (1969), "Preference and Perception Measures in New Product Development: An Exposition and Review," *Industrial Management Review,* 11 (Fall), 21–37.

Silk, A.J., and G.L. Urban (1978), "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research,* 15 (May), 171–191.

Tauber, E.M. (1977), "Forecasting Sales Prior to Test Market," *Journal of Marketing,* 41 (January), 80–84.

Urban, G.L. (1970), "SPRINTER Mod III: A Model for the Analysis of New Frequently Purchased Consumer Products," *Operations Research,* 18 (September), 805–855.

von Hippel, E. (1978), "Successful Industrial Products from Customer Ideas," *Journal of Marketing,* 42 (January), 39–49.

Wind, Y. (1974), "A Note on the Classification and Evaluation of New Product Forecasting Models," presented at the American Marketing Association Conference, April.

Wind, Y. and H.J. Claycamp (1976), "Planning Product Line Strategy: A Matrix Approach," *Journal of Marketing,* 40 (January), 2–9.

Appendix 18A: Customer Flow and Rate of Customer Flow Equations

The customer flow equations as presented in the text are,

$$S_1(t) + S_2(t) + N(t) = M(t) \qquad (18A.1)$$

where

$$S_1(t) = M(t) - \bar{N}(t) \qquad (18A.2)$$

$$S_2(t) = \bar{N}(t) - N(t) \qquad (18A.3)$$

The rate of the flow of customers across the three segments can be obtained by differentiating the above three equations, i.e.,

$$\frac{dS_1(t)}{dt} = \frac{dM(t)}{dt} - \frac{d\bar{N}(t)}{dt} \qquad (18A.4)$$

$$\frac{dS_2(t)}{dt} = \frac{d\bar{N}(t)}{dt} - \frac{dN(t)}{dt} \qquad (18A.5)$$

and

$$\frac{dN(t)}{dt} = \frac{dM(t)}{dt} - \frac{dS_1(t)}{dt} - \frac{dS_2(t)}{dt} \qquad (18A.6)$$

In the discrete-time representation, rate equations are

$$s_1(t) = m(t) - \bar{n}(t) \qquad (18A.7)$$

$$s_2(t) = \bar{n}(t) - n(t) \qquad (18A.8)$$

$$n(t) = m(t) - s_1(t) - s_2(t) \qquad (18A.9)$$

where

$s_1(t) = S_1(t + 1) - S_1(t)$; $s_2(t) = S_2(t + 1) - S_2(t)$;

$m(t) = M(t + 1) - M(t)$; $\bar{n}(t) = \bar{N}(t + 1) - \bar{N}(t)$; and

$n(t) = N(t + 1) - N(t)$.

The customer flow equations, 18A.1 to 18A.3, and the rate equations

18A.4 to 18A.6 or 18A.7 to 18A.9, form the basis for all the proposed diffusion models.

The basic diffusion models (Bass 1969; Fourt and Woodlock 1960; Mansfield 1961; and the Gompertz curve) assume a constant total population of the potential customers; i.e., $\bar{N}(t)$ is constant or $\frac{d\bar{N}(t)}{dt} = 0$. Recalling equations 18A.1 to 18A.9, these models consider the following customer and rate equations:

Customer flow:

$$S_2(t) = \bar{N} - N(t) \tag{18A.10}$$

Rate of flow:

$$\frac{dN(t)}{dt} = -\frac{dS_2(t)}{dt} \tag{18A.11}$$

or

$$n(t) = -s_2(t) \tag{18A.12}$$

Note in equations 18A.11 and 18A.12 that the decrease in the number of potential customers at any time t is equal to the simultaneous increase in the number of adopters at time t. Because of this relationship, the *basic* models are *primarily concerned with modeling $n(t)$ or $N(t)$*. Equations 18A.10 to 18A.12 also represent the customer flow and rate of customer flow for the extensions to the basic models proposed by Bass (1978), Horsky and Simon (1978), Lilien and Rao (1978), Peterson and Mahajan (1978), and Robinson and Lakhani (1975).

Modeling the product growth by considering the customer flow across all three segments of the diffusion process, the extensions proposed by Chow (1967), Dodson and Muller (1978), Lackman (1978), and Mahajan and Peterson (1978) implicitly assume that $M(t)$—the number of individuals in the total market—is constant. Recalling equations 18A.1 to 18A.9, these models consider the following customer flow and rate equations:

Customer flow:

$$S_1(t) = M - \bar{N}(t) \tag{18A.13}$$

$$S_2(t) = \bar{N}(t) - N(t) \tag{18A.14}$$

$$N(t) = M - S_1(t) - S_2(t) \tag{18A.15}$$

Rate of flow:

$$-\frac{dS_1(t)}{dt} = \frac{d\bar{N}(t)}{dt} \tag{18A.16}$$

$$\frac{dS_2(t)}{dt} = \frac{d\bar{N}(t)}{dt} - \frac{dN(t)}{dt} \tag{18A.17}$$

$$\frac{dN(t)}{dt} = -\frac{dS_1(t)}{dt} - \frac{dS_2(t)}{dt} \tag{18A.18}$$

Equation 18A.17 implies that the increase in number of potential customers at any time t is a net result of the decrease in the untapped market (individuals who have transferred from the untapped market to the potential market) and increase in the number of adopters (individuals who have transferred from the potential market to the current market). Because of this relationship, these *models are concerned with modeling* $n(t)$ and $\bar{n}(t)$. If $n(t)$ and $\bar{n}(t)$ are known, the flow and the rate of flow of customers across the three segments is completely specified.

In extending the basic diffusion models to include the third segment, untapped market, Mahajan et al (1979) relax the assumption that $M(t)$—the number of individuals in the market—is constant. Equations 18A.1 to 18A.9 specify the customer flow and rate of flow for their model. In addition to modeling $n(t)$, $\bar{n}(t)$, *this model is also concerned with modeling* $m(t)$. The knowledge of $n(t)$, $\bar{n}(t)$, and $m(t)$ completely specifies the customer flow across the segments.

References

Bass, F.M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science,* 15 (January) 215–227.

——— (1978), "An Integration of the New Product Growth Model with the Experience Cost Function and Optimal Pricing," presented at the TIM/ORSA National Meeting in New York City, May.

Chow, G.C. (1967), "Technological Change and the Demand for Computers," *American Economic Review,* 57 (December), 1117–1130.

Dodson, J.A., and E. Muller (1978), "Models of New Product Diffusion through Advertising and Word-of-Mouth," *Management Science,* 15 (November), 1568–1578.

Fourt, L.A., and J.W. Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* 25 (October), 31–38.

Horsky, D., and L.S. Simon (1978), "Advertising in a Model of New Product Diffusion," presented at the TIMS/ORSA National Meeting in New York City, May.

Lackman, C.L. (1978), "Gompertz Curve Forecasting: A New Product Application," *Journal of the Market Research Society,* 20 (January), 45-47.

Lilien, G.L., and A.G. Rao (1978), "A Marketing Promotion Model with Word of Mouth Effect," Working Paper 976-78, Cambridge, Mass.: Sloan School of Management, Massachusetts Institute of Technology.

Mahajan, V., and R.A. Peterson (1978), "Innovation Diffusion in a Dynamic Potential Adopter Population," *Management Science,* 15 (November), 1589-1597.

Mahajan, V., R.A. Peterson, A.K. Jain, and N. Malhotra (1979), "A New Product Growth Model with a Dynamic Market Potential," *Long Planning* (forthcoming).

Mansfield, E. (1961), "Technical Change and the Rate of Imitation," *Econometrica,* 29 (October), 741-766.

Peterson, R.A., and V. Mahajan (1978), "Multi-Product Growth Models," in *Research in Marketing,* J. Sheth, ed., Greenwich, Conn.: JAI Press.

Appendix 18B

Incorporating the effect of advertising into the growth model and considering the customer flow across all the three segments of the diffusion process, Dodson and Muller (1978) proposed the following rate equations for durables:

$$-\frac{dS_1(t)}{dt} = k_1S_1(t)\,(S_2(t) + N(t)) + k_2S_1(t) \tag{18B.1}$$

$$\frac{dS_2(t)}{dt} = k_1S_1(t)(S_2(t) + N(t)) + k_2S_1(t) - aS_2(t) \tag{18B.2}$$

$$\frac{dN(t)}{dt} = aS_2(t) \tag{18B.3}$$

where equation 18B.1 represents the decrease in the number of individuals in the untapped market (i.e., individuals who have transferred from the untapped market segment to the potential market), equation 18B.2 represents the net increase in the number of potential customers (number of individuals who have transferred from the untapped market to the potential market minus the number of individuals who have transferred from the potential market to the current market) and equation 18B.3 represents the increase in the number of adopters. Note that in their model

$$\frac{dS_1(t)}{dt} + \frac{dS_2(t)}{dt} + \frac{dN(t)}{dt} = 0 \tag{18B.4}$$

That is, recalling the general rate equation 18A.6 for the flow of customers across the three market segments, it is clear that Dodson and Muller assume $dM(t)/dt = 0$, i.e., the total number of individuals in the market, $M(t)$, is constant. Second, because of this relatioship, the customer flow across the three segments is completely specified by *two* rate equations. That is, if $dS_1(t)/dt$ and $dN(t)/dt$ are known,

$$\frac{dS_2(t)}{dt} = -\frac{dS_1(t)}{dt} - \frac{dN(t)}{dt} \tag{18B.5}$$

Now, from equations 18A.13, 18A.14, and 18A.16, since $S_1(t) = M -$

$\bar{N}(t)$, $S_2(t) = \bar{N}(t) - N(t)$, and $-dS_1(t)/dt = d\bar{N}(t)/dt$, the model proposed by Dodson and Muller, equations 18B.1 and 18B.3, reduces to the following rate equations:

$$\frac{d\bar{N}(t)}{dt} = k_1\bar{N}(t)(M - \bar{N}(t) + k_2(M - \bar{N}(t) \qquad (18B.6)$$

$$\frac{dN(t)}{dt} = a(\bar{N}(t) - N(t)$$

The discrete formulation of equation 18B.6 is given by equation 18.26.

Reference

Dodson, J.A., and E. Muller (1978), "Models of New Product Diffusion through Advertising and Word of Mouth," *Management Science,* 15 (November), 1568–1578.

19 A New-Product Growth Model for Consumer Durables

Frank M. Bass

The concern of this paper is the development of timing of initial purchase of new consumer products. The empirical aspects of the work presented here deal exclusively with consumer durables. The theory, however, is intended to apply to the growth of initial purchases of a broad range of distinctive "new" generic classes of products. Thus, we draw a distinction between new classes of products as opposed to new brands or new models of older products. While further research concerning growth rate behavior is currently in process for a wider group of products, attention focuses here exclusively upon infrequently purchased products.

Haines [6], Fourt and Woodlock [5], and others have suggested growth models for new brands or new products which suggests exponential growth to some asymptote. The growth model postulated here, however, is best reflected by growth patterns similar to that shown in figure 19-1. Sales grow to a peak and then level of at some magnitude lower than the peak. The stabilizing effect is accounted for by the relative growth of the replacement purchasing component of sales and the decline of the initial purchase component. We shall be concerned here only with the timing of initial purchase.

Long-range forecasting of new product sales is a guessing game, at best. Some things, however, may be easier to guess than others. The theoretical framework presented here provides a rationale for long-range forecasting. The theory stems mathematically from the contagion models which have found such widespread application in epidemology [2]. Behaviorally, the assumptions are similar in certain respects to the theoretical concepts emerging in the literature on new product adoption and diffusion [7, 8, 9, 13], as well as to some learning models [3, 12]. The model differs from models based on the log-normal distribution [1, 4, 10] and other growth models [11] in that the behavioral assumptions are explicit.

Some of the basic ideas in this paper were originally suggested to the author by Peter Frevert, now of the University of Kansas. Thomas H. Bruhn, Gordon Constable, and Murray Silverman provided programming and computational assistance. Reprinted from F.M. Bass, "A New Product Growth Model for Consumer Durables," *Management Science* 15 (January 1969):215–227, copyright 1969. The Institute of Management Sciences.

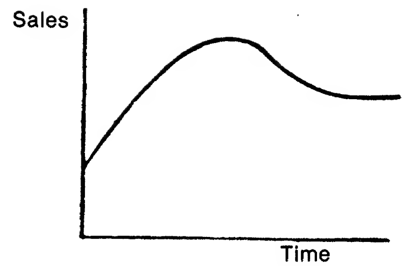


Figure 19–1. Growth of a New Product

The Theory of Adoption and Diffusion

The theory of the adoption and diffusion of new ideas or new products by a social system has been discussed at length by Rogers [13]. This discussion is largely literary. It is, therefore, not always easy to separate the premises of the theory from the conclusions. In the discussion which follows an attempt will be made to outline the major ideas of the theory as they apply to the *timing* of adoption.

Some individuals decide to adopt an innovation independently of the decisions of other individuals in a social system. We shall refer to these individuals as *innovators*. We might ordinarily expect the first adopters to be innovators. In the literature, the following classes of adopters are specified: (1) Innovators; (2) Early Adopters; (3) Early Majority; (4) Late Majority; and (5) Laggards. This classification is based upon the timing of adoption by the various groups.

Apart from innovators, adopters are influenced in the timing of adoption by the pressures of the social system, the pressure increasing for later adopters with the number of previus adopters. In the mathematical formulation of the theory presented here we shall aggregate groups (2) through (5) above and define them as *imitators*. Imitators, unlike innovators, are influenced in the timing of adoption by the decisions of other members of the social system. Rogers defines innovators, rather arbitrarily, as the first two and one-half percent of the adopters. Innovators are described as being venturesome and daring. They also interact with other innovators. When we say that they are not influenced in the timing of purchase by other members of the social system, we mean that the pressure for adoption, for this group, does not increase with the growth of the adoption process. In fact, quite the opposite may be true.

In applying the theory to the timing of initial purchase of a new consumer product, we formulate the following precise and basic assumption which, hopefully, characterizes the literary theory: *The probability that an*

initial purchase will be made at T given that no purchase has yet been made is a linear function of the number of previous buyers. Thus, $P(T) = p + (q/m)Y(T)$, where p and q/m are constants and $Y(T)$ is the number of previous buyers. Since $Y(0) = 0$, the constant p is the probability of an initial purchase at $T = 0$ and its magnitude reflects the importance of innovators in the social system. Since the parameters of the model depend upon the scale used to measure time, it is possible to select a unit of measure for time such that p reflects the fraction of all adopters who are innovators in the sense in which Rogers defines them. The product q/m times $Y(T)$ reflects the pressures operating on imitators as the number of previous buyers increases.

In the section which follows, the basic assumption of the theory will be formulated in terms of a continuous model and a density function of time to initial purchase. We shall therefore refer to the linear probability element as a likelihood.

Assumptions and the Model

The following assumptions characterize the model:

a) Over the period of interest ("life of the product") there will be m initial purchases of the product. Since we are dealing with infrequently purchased products, the unit sales of the product will coincide with the number of initial purchases during that part of the time interval for which replacement sales are excluded. After replacement purchasing begins, sales will be composed of both initial purchases and replacement purchases. We shall restrict our interest in sales to that time interval for which replacement sales are excluded, although our interest in initial purchase will extend beyond this interval.

b) The likelihood of purchase at time T given that no purchase has yet been made is

$$[f(T)]/[1 - F(T)] = P(T) = p + q/mY(T) = p + qF(T),$$

where $f(T)$ is the likelihood of purchase at T and

$$F(T) = \int_0^T f(t)\,dt, \qquad F(0) = 0$$

Since $f(T)$ is the likelihood of purchase at T and m is the total number purchasing during the period for which the density function was constructed,

$$Y(T) = \int_0^T S(t)dt = m \int_0^T f(t)dt = mF(t)$$

is the total number purchasing in the $(0, T)$ interval. Therefore, sales at $T =$

$$S(T) = mf(T) = P(T)[m - Y(T)]$$

$$= \left[p + q\int_0^T S(t)dt/m\right]\left[m - \int_0^T S(t)dt\right].$$

Expanding this product we have

$$S(T) = pm + (q - p)Y(T) - q/m[Y(T)]^2.$$

The behavioral rationale for these assumptions is summarized:

1. Initial purchases of the product are made by *both* "innovators" and "imitators," the important distinction between an innovator and an imitator being the buying influence. Innovators are not influenced in the timing of their initial purchase by the number of people who have already bought the product, while imitators are influenced by the number of previous buyers. Imitators "learn," in some sense, from those who have already bought.
2. The importance of innovators will be greater at first but will diminish monotonically with time.
3. We shall refer to p as the coefficient of innovation and q as the coefficient of imitation.

Since

$$f(T) = [p + qF(T)][1 - F(T)] = p + (q - p)F(T) - q[F(T)]^2,$$

in order to find $F(T)$ we must solve this non-linear differential equation:

$$dT = dF/(p + (q - p)F - qF^2).$$

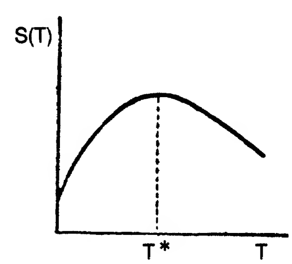


Figure 19-2. Growth Rate ($q > p$)

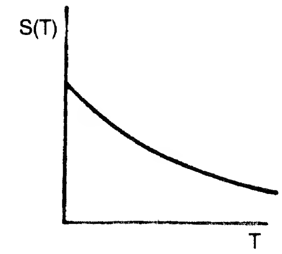


Figure 19-3. Growth Rate ($q \leq p$)

The solution is:

$$F = (q - pe^{-(T+C)(p+q)})/q(1 + e^{-(T+C)(p+q)}).$$

Since $F(0) = 0$, the integration constant may be evaluated:

$-C = (1/(p + q))\mathrm{Ln}(q/p)$ and $F(T) = (1 - e^{-(p+q)T})/(q/pe^{-(p+q)T} + 1)$

Then,

$$f(T) = ((p + q)^2/p)[e^{-(p+q)T}/(q/pe^{-(p+q)T} + 1)^2],$$

and

$$S(T) = (m(p + q)^2/p)[e^{-(p+q)T}/(q/pe^{-(p+q)T} + 1)^2].$$

To find the time at which the sales rate reaches its peak, we differentiate S,

$$S' = (m/p(p + q)^3e^{-(p+q)T}(q/pe^{-(p+q)T} - 1))/(q/pe^{-(p+q)T} + 1)^3$$

Thus, $T^* = -1/(p + q)\,\mathrm{Ln}\,(p/q) = 1/(p + q)\,\mathrm{Ln}\,(q/p)$ and if an interior maximum exists, $q > p$. The solution is depicted graphically in figure 19-2 and 19-3. We note that $S(T^*) = (m(p + q)^2/4q$ and $Y(T^*) = \int_0^{T^*} S(t)dt = m(q - p)/2q$. Since for successful new products the coefficient of imitation will ordinarily be much larger than the coefficient of innovation, sales will attain its maximum value at about the time that cumulative sales are approximately one-half m. We note also that the expected time to purchase, (E/T), is $1/q\ \mathrm{Ln}\ ((p + q)/p)$.

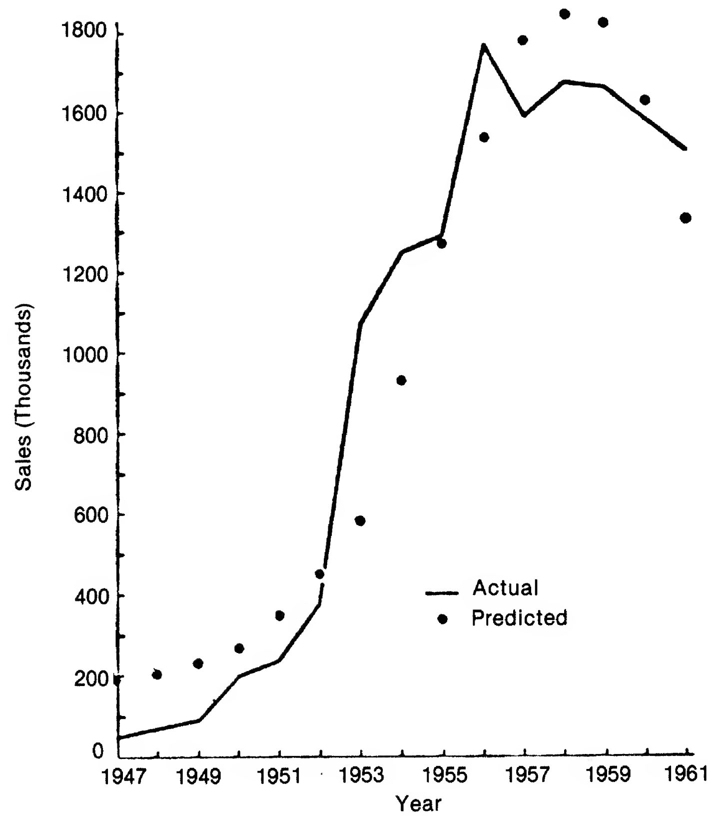


Figure 19-4. Actual Sales and Sales Predicted by Regression Equation

The Discrete Analogue

The basic model is: $S(T) = pm + (q - p)Y(T) - q/mY^2(T)$. In estimating the parameters, p, q and m from discrete time series data we use the following analogue: $S_T = a + bY_{T-1} + cY^2_{T-1}$. $T = 2, 3 \ldots$ where: S_T = sales at T, and $Y_{T-1} = \Sigma_{t=1}^{T-1} S_t$ = cumulative sales through period $T - 1$. Since a estimates pm, b estimates $q - p$, and c estimates $-q/m$: $-mc = q$, $a/m = p$.

Then $q - p = -mc - a/m = b$, and $cm^2 + bm + a = 0$, or $m = (-b \pm \sqrt{b^2 - 4ca}/2c)$ and parameters p, q, and m are identified. If we write $S(Y_{T-1})$ and differentiate with respect to Y_{T-1}, $\frac{dS_T}{dY_{T-1}} = b + 2cY_{T-1}$. Setting this equal to 0 $(Y^*_{T-1}) = -b/2c = m(q - p)/2q = Y(T^*)$, and $S_T(Y^*_{T-1}) = a - b^2/2c + b^2/4c = m(p + q)^2/4q = S(T^*)$. Therefore, the

Table 19–1
Growth Model Regression Results for Eleven Consumer Durable Products

Product	*Period Covered*	$a(10^2)$	b	$c(10^{-7})$	R^2	a/σ_a	b/σ_b	c/σ_c	$m(10^3)$	p	q
Electric refrigerators	1920–1940	104.67	0.21305	−0.053913	0.903	1.164	6.142	−2.548	40,001	0.0025167	0.21566
Home freezers	1946–1961	308.12	0.15298	−0.077868	0.742	4.195	4.769	−3.619	21,973	0.018119	0.17110
Black-and-white television	1946–1961	2,696.2	0.22317	−0.025957	0.576	3.312	3.724	−3.167	96,717	0.027877	0.25105
Water softeners	1949–1961	0.10256	0.27925	−512.59	0.919	3.593	8.089	−6.451	5,793	0.017703	0.29695
Room air conditioners	1946–1961	175.69	0.40820	−0.24777	0.911	1.915	8.317	−6.034	16,895	0.010299	0.41861
Clothes dryers	1948–1961	259.67	0.33968	−0.23647	0.896	2.941	7.427	−5.701	15,092	0.017206	0.35688
Power lawnmowers	1948–1961	410.98	0.32871	−0.075506	0.932	1.935	7.408	−4.740	44,751	0.0091837	0.33790
Electric bed coverings	1949–1961	450.04	0.23800	−0.031842	0.976	3.522	6.820	−1.826	76,589	0.005876	0.24387
Automatic coffee makers	1948–1961	1,008.2	0.28435	−0.051242	0.883	3.109	6.186	−4.353	58,838	0.017135	0.30145
Steam irons	1949–1960	1,594.7	0.29928	−0.058875	0.828	3.649	5.288	−4.318	55,696	0.028632	0.32791
Recover players	1952–1961	543.94	0.62931	−0.29817	0.899	1.911	5.194	−3.718	21,937	0.024796	0.65410

Data Sources: *Economic Almanac, Statistical Abstracts of the U.S. Electrical Merchandising,* and *Electric Merchandising Week.*

maximum value of S as a function of time coincides with the maximum value of S as a function of cumulative sales.

Regression Analysis

In order to test the model, regression estimates of the parameters were developed using annual time series data for eleven different consumer durables. The period of analysis was restricted in every case to include only those intervals in which repeat purchasing was not a factor of importance. These intervals were determined on the basis of a subjective appraisal of the durability of the product as well as from limited published data concerning "scrappage rates" and repurchase cycles.

Table 19–1 displays the regression results. The data appear to be in good agreement with the model. The R^2 values indicate that the model describes the growth rate behavior rather well. Furthermore, the parameter estimates seem reasonable for the model. The regression estimates for the parameter c are negative in every case, as required in order for the model to make sense, and the estimates of m are quite plausible. One of the more important contributions derived from the regression analysis is the implied estimate of the total number of initial purchases to be made over the life of the product. Figures 19–4, 19–5, and 19–6 show the actual values of sales and the values predicted by the regression equation for three of the products analyzed. For every product studied the regression equation describes the general trend of the time path of growth very well. In addition, the regression equation provides a very good fit with respect to both the magnitude and the timing of the peaks for all of the products. Deviations from trend are largely explainable in terms of short-term income variations. This is especially apparent in figure 19–5, where it is easy to identify recessions and booms in the years of sharp deviations from trend.

Model Performance

The performance of the regression equation relative to actual sales is a relatively weak test of the model's performance since it amounts to an *ex post* comparison of the regression equation estimates with the data. A much stronger test is the performance of the basic model with time as the variable and controlling parameter values as determined from the regression estimates. Table 19–2 provides a comparison of the model's prediction of time of peak and magnitude of peak for the eleven products studied.

Since according to the model $S(0) = pm$, we identify time period 1 as that period in which sales equal or exceed pm for the first time. It is clear

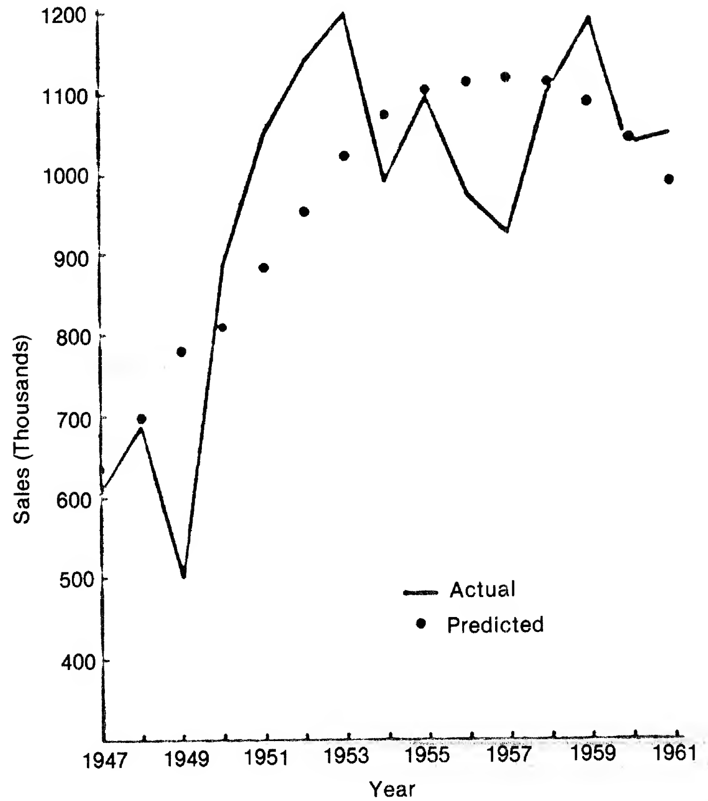


Figure 19–5. Actual Sales and Sales Predicted by Regression Equation (Home Freezer)

from the comparison shown in table 19–2 that the model provides good predictions of the timing and magnitude of the peaks for all eleven products studied.

In order to determine the accuracy with which it would have been possible to "forecast" period sales over a long-range interval with prior knowledge of the parameter values, the regression estimates of the parameters were substituted in the basic model, $S(T) = [m(p + q)^2/p][e^{-(p+q)T}/(q/pe^{-(p+q)T} + 1)^2]$, and sales estimates generated for each of the products for each year indicated in the intervals shown in table 19–3. In most cases the model provides a good fit to the data. Even in the few instances of low r^2 values, the model provides a good description of the general trend of the sales curve, the deviations from trend being sharp, but ephemeral. Figures

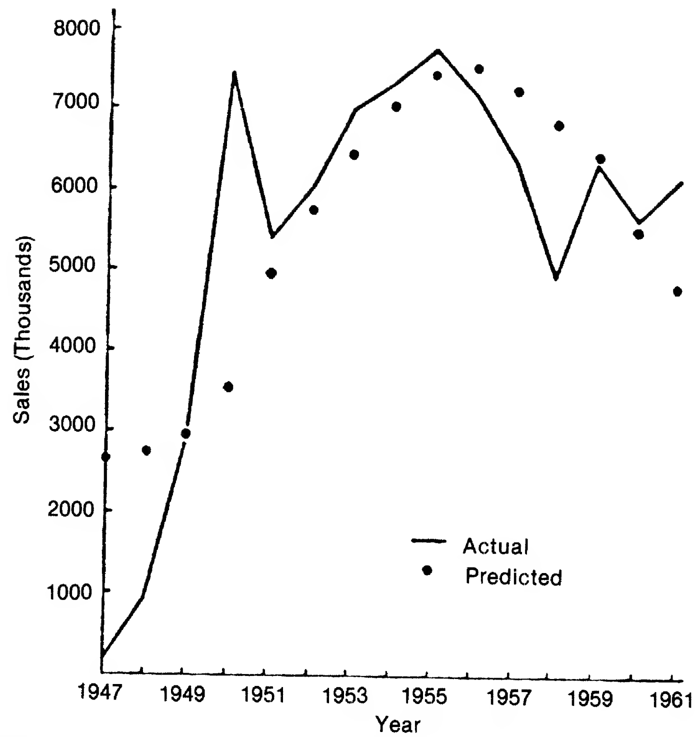


Figure 19–6. Actual Sales and Sales Predicted by Regression Equation (Black-and-White Television)

19–7, 19–8, and 19–9 illustrate the predicted and actual sales curves for three of the products.

It would appear fair to conclude that the data are in generally good agreement with the model. The model has, then, in some sense, been "tested" and verified. We may now claim to know something about the phenomenon we set out to explore. The questions is, however, will this knowledge be useful for purposes of long-range forecasting?

Long-Range Forecasting

There are two cases worth considering in long-range forecasting: the no-data case and the limited-data case. For either of these possibilities one may well ask: Is it easier to guess the sales curve for the new product or easier to guess the parameters of the model? No attempt will be made here to answer

Table 19-2
Comparison of Predicted Time and Magnitude of Peak with Actual Values for Eleven Consumer Durable Products

Product	q/p	*Predicted Time of Peak* $T^* = 1/(p+q)) \ln(q/p)$	*Actual Time of Peak*[a]	*Predicted Magnitude* $S(T^*) = m(p+q)^3/(4q)$ (10^4)	*Actual Magnitude of Peak* (10^4)
Electric refrigerators	82.4	20.1	[b]	2.20	[b]
Home freezers	9.4	11.6	13	1.2	1.2
Black-and-white television	9.0	7.8	7	7.5	7.8
Water softeners	16.7	8.9	9	0.5	0.5
Room air conditioners	40.2	8.6	7	1.8	1.7
Clothes dryers	20.7	8.1	7	1.5	1.5
Power lawnmowers	36.7	10.3	11	4.0	4.2
Electric bed coverings	41.6	14.9	14	4.8	4.5
Automatic coffee makers	18.1	9.0	10	4.8	4.9
Steam irons	11.4	6.8	7	5.5	5.9
Recover players	26.3	4.8	5	3.8	3.7

[a]Time period one is defined as that for which sales equal or exceed *pm* for the first time.
[b]Interrupted by war. Prewar peak in year 16 (1940) at 2.6×10^6 units.

Table 19-3
Forecasting Accuracy of the Model for Eleven Consumer Durable Products

Product	*Period of Forecast*	r^2
Electric refrigerators	1926–1940	.762
Home freezers	1947–1961	.473
Black-and-white television	1949–1961	.077[a]
Water softeners	1950–1961	.920
Room air conditioners	1950–1961	.900
Clothes dryer	1950–1961	.858
Power lawnmowers	1949–1961	.898
Electric bed coverings	1950–1961	.934
Automatic coffee makers	1951–1961	.690
Steam irons	1950–1961	.730
Recover players	1953–1958	.953

[a]The low "explained" variance for this product is accounted for by extreme deviation from trend in two periods. Actually, the model provides a fairly good description of the growth rate, as indicated in figure 19–9.

this question, in general, but it does seem likely that for some products it would be possible to make plausible guesses of the parameters. Analysis of the potential market and the buying motives should make it possible to guess at m, the size of the market, and of the relative values of p and q, the latter guess being determined by a consideration of buying motives. If the sales curve is to be determined by means other than the model suggested in this paper, the implications of this forecast in terms of the parameters of the model might be useful as a test of the credibility of the forecast.

In order to illustrate the forecasting possibilities in the limited data case, we shall develop a forecast for color television set sales. In principle, since there are three parameters to be estimated, some kind of estimate is possible with only three observations, and the first of these observations occurs at $T = 0$. Any such estimate should be viewed with some skepticism, however, since the parameter estimates are very sensitive to small variations in the three observations. Before applying estimates obtained from a limited number of observations, the plausibility of these estimates should be closely scrutinized.

In substituting $\Sigma_{t+0}^{T-1} S_t$ in the descrete analogue for $\int_0^T S(t)dt$ in the continuous model, a certain bias was introduced. This bias is mitigated when

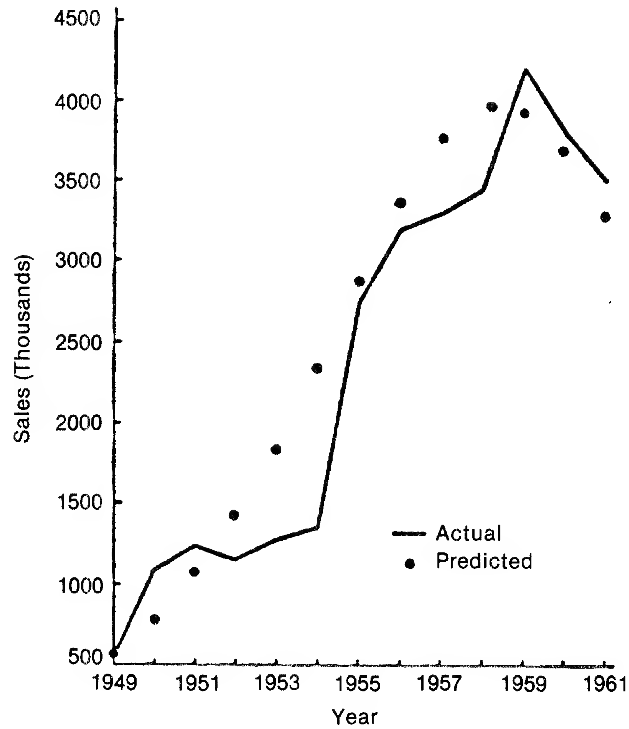


Figure 19-7. Actual Sales and Sales Predicted by Model (Power Lawnmowers)

there are several observations, but can be crucial when there are only a few. Thus, the proper formulation of the discrete model, if $S_T = S(T)$: is $S_T = a + bk(T)Y_{T-1} + ck^2(T)Y^2_{T-1}$, where $k(T)/Y_{T-1}$. We note that for any probability distribution for which: a) $f(x) = 1/k\,[F(x-1) - F(x)]$, and b) $F(0) = 0$,

$\Sigma_{t=0}^{T-1} f(t) = 1/kF(x)$. In particular, these two properties hold for the exponential distribution. Therefore, for this distribution $\Sigma_{t=0}^{T-1} F(x)/f(t) = k$. The density function $f(T)$ in the growth model developed in this paper is approximately exponential in character *when p and T are small.* Thus, $f_{epx}(T) = 1/k\,[F_{epx}(T+1) - F_{epx}(T)]$ and $1/k = (p+q)/[e^{(p+q)} - 1]$. For small values of T we therefore write: $S_T = a + b'Y_{T-1} + c'Y^2_{T-1}$, where $b' = kb$, and $c' = k^2c$. Then $m = km'$, $q = 1/kq'$, and $p = 1/kp'$. The value of $1/k$ for each of several different values of $p + q$ has been calculated and appears in table 19-4.

Table 19-4
Calcualted Values of $1/k$ and $(p+q)$

$p+q$	$1/k = (p+q)/[e^{p+q}-1]$
0.3	.85
0.4	.81
0.5	.77
0.6	.73
0.7	.69
0.8	.65
0.9	.61

On the basis of the relationship between k and $(p + q)$ indicated in table 19-4: $1/k = .97 - .4(p + q)$, $(p'/q')q = q = 1/kq' = .97q'/(1 + .4(1 + 1/\theta)q')$, where $\theta = q'/p' = q/p$, and $p = (p'/q')q' = .97p'/(1 + .4(1 + \theta)p')$.

We turn now to the forecast of color television set sales. The following data are available:

Sales (Millions of Units)	*Year*
.7	1963
1.35	1964
2.50	1965

Solving the following system of equations:

$$S_0 = .7 = a$$

$$S_1 = 1.35 = a + .7b' + .49c'$$

$$S_2 = 2.50 = a + 2.05b' + 4.20c'$$

we find

$$a' = .7 \qquad b' = .954 \qquad c' = -.0374$$

$$m' = 26.2 \qquad q' = .96 \qquad p' = .0267$$

$$q = .67 \qquad p = .018 \qquad m = 37.4$$

Since these parameter values appear plausible, they have been used in the basic model to generate the series of estimates of sales shown in table 19-5

Table 19-5
Forecast of Color-Television Sales 1966-1970

Year	*Sales (Millions)*
1964	1.35
1965	2.5
1966	4.1
1967	5.8
1968	6.7
1969	6.3
1970	4.7

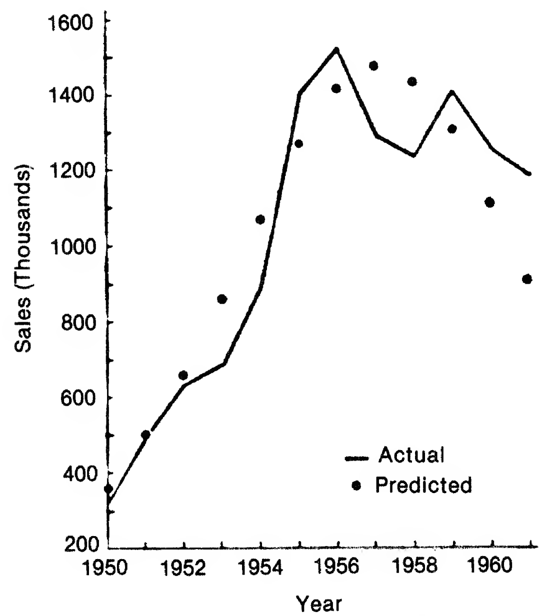


Figure 19-8. Actual Sales and Sales Predicted by Model (Clothes Dryers)

and figure 19-10. The projected peak occurs in 1963 at around 7 million units. This forecast differs somewhat from some industry forecasts. At this writing, one company's research department has estimated that sales will "top out" in 1967 at between 7 and 8 million units. The forecast speaks for itself, and the ultimate reality of actual sales and one's personal criterion of "goodness" will determine whether or not the forecast was a good one.

The preceding forecast was made in late 1966. The following report was published in the May 19, 1967 issue of the *Wall Street Journal:*

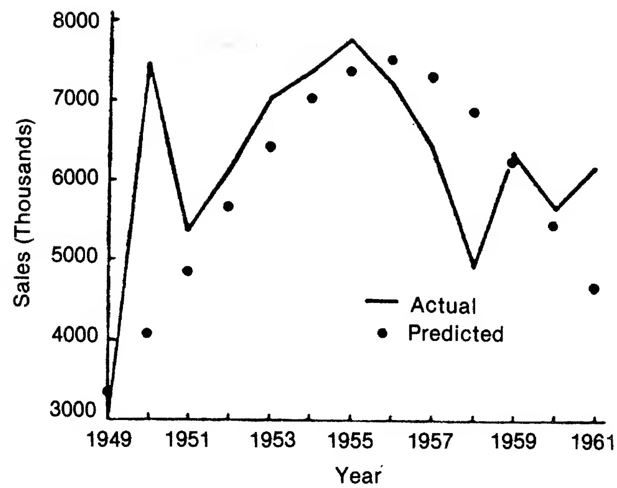


Figure 19-9. Actual Sales and Sales Predicted by Model (Black-and-White Television).

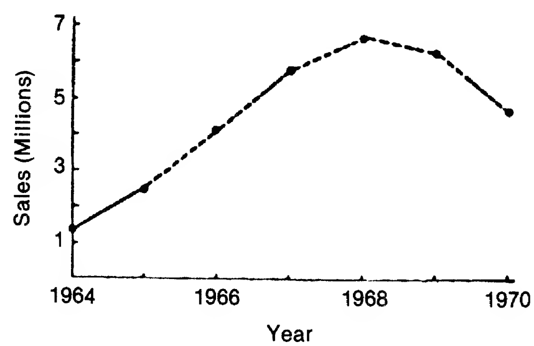


Figure 19-10. Projected Sales—Color Television

> . . . Industry sources say most color set producers are continuing to trim production. "Only R.C.A., Zenith Radio Corp. and Philco-Ford Corp. aren't cutting back now, and they're taking a big gamble by banking on a big sales pickup this fall," asserts one industry analyst.
>
> Executives at R.C.A., the biggest producer of color sets, have turned reticent about publicly forecasting industry sales for 1967. Their silence contrasts markedly with last fall, when R.C.A. and other companies were exuberantly predicting industry sales of 7 million color sets in 1967.
>
> Admiral recently pared it's 1967 forecast to 6.1 million color-set sales for the industry. . . .

If the model developed in this paper does nothing else, it does demonstrate vividly the slowing down of growth rates as sales near the peak. In focusing upon the vital theoretical issues the model may serve to aid management in avoiding some of the more obviously absurd forecasts as have been made in the past.

While our forecast of color television sales was objectively determined in the sense that it was derived from data, it is also based upon a subjective judgment of the plausibility of the parameters. Since the parameter estimates are very sensitive to small variations in the observations when there are only a few observations, the importance of the plausibility test cannot be overemphasized. The parameter for which one has the strongest intuitive feeling is m. The plausibility test for this parameter is therefore perhaps the most easily derived. An examination of the parameter values implied by early sales of the other products analyzed in this paper indicates that in some instances these are remarkably close to the regression estimates, while in others they differ markedly. In every instance in which the early year estimates differ substantially from the regression estimates, these estimates are easily rejected on the basis of the implausibility of the implied value of m.

Conclusion

The growth model developed in this paper for the timing of initial purchase of new products is based upon an assumption that the probability of purchase at any time is related linearly to the number of previous buyers. There is a behavioral rationale for this assumption. The model implies exponential growth of initial purchases to a peak and then exponential decay. In this respect it differs from other new product growth models.

Data for consumer durables are in good agreement with the model. Parameter estimates derived from regression analysis when used in conjunction with the model provide good descriptions of the growth of sales. From a planning viewpoint, probably the central interest in long-range forecasting lies in predictions of the timing and magnitude of the sales peak. The model provides good predictions of both of these variables for the products to which it has been applied. Insofar as the model contributes to an understanding of the process of new product adoption, the model may be useful in providing a rationale for long-range forecasting.

References

1. Bain, A.D., *The Growth of Television Ownership in the United Kingdom, A Lognormal Model,* Cambridge, The University Press, 1964.

2. Bartlett, M.S., *Stochastic Population Models in Ecology and Epidemiology,* London: Methuen, 1960.
3. Bush, R.R., and Mosteller, F., *Stochastic Models for Learning,* New York: Wiley, 1955.
4. Dernburg, T.F., "Consumer Response to Innovation: Television," in *Studies in Household Economic Behavior,* Yale Studies in Economics, Vol. 9, New Haven, Connecticut: Yale, 1955.
5. Fourt, L.A., and Woodlock, J.W., "Early Prediction of Market Success for New Grocery Products," *Journal of Marketing,* No. 2, Vol. 26, October, 1960.
6. Haines, G.H., Jr., "A Theory of Market Behavior after Innovation, *Management Science,* No. 4, Vol. 10, July, 1964.
7. Katz, E., and Lazarsfeld, F., *Personal Influence,* New York: The Free Press, 1955.
8. King, C.W., "Adoption and Diffusion Research in Marketing: An Overview," in *Science, Technology and Marketing,* 1966 Fall Conference of the American Marketing Association, Chicago, 1966.
9. Mansfield, E., "Technological Change and the Rate of Imitation," *Econometrica,* No. 4, Vol. 29, October, 1961.
10. Masst, W.F., "Innovation and Market Penetration," Ph.D. Thesis, Massachusetts Institute of Technology, 1960, Cambridge, Massachusetts.
11. Pessemier, E.A., *New Product Decisions, An Analytic Approach,* New York: McGraw-Hill, 1966.
12. Rashevsky, N., *Mathematical Biology of Social Behavior,* Chicago: The University Press, 1959.
13. Rogers, E.M., *Diffusion of Innovations,* New York: The Free Press, 1962.

20 A Simple Mathematical Theory of Innovative Behavior

David F. Midgley

Introduction

The quantitative theory of innovative behavior presented in the following pages attempts to describe the adoption or first purchase of an innovation and represents a first step in the formulation of a general theory of the product life cycle. While the theory is firmly based on the known psychology and sociology of the consumer, it is *mathematically* stated. It becomes possible, therefore, to test this formulation more rigorously than the "traditional" schema proposed by Rogers (1962). In the cases to be discussed here the theory has been applied to the sales of new products in the mass market, to low-risk, fast-moving, nondurable products such as toothpastes, packaged foodstuffs and detergents. However, the theory itself is perfectly capable of being generalized to a wider range of situations—in fact, to any situation where a new product of idea is diffused throughout a society by the means of interpersonal communication, with or without other communication channels, and where it is possible for individuals to form attitudes toward the innovation which are favorable, unfavorable, or neutral. Low-risk products were chosen for the research, because in the United Kingdom the only reliable data detailing adoption curves relate to this type of product and are obtained by the use of consumer panels.

The more obvious origins of the research lie in the field of the diffusion of innovations, a body of knowledge which has been admirably collated and synthesized by Rogers and Shoemaker (1971). It is assumed that by now most readers will be familiar with the basic concepts of this academic research tradition—the adoption curve, adopter categories, innovativeness, and opinion leadership. There is only enough space here to deal with aspects that impinge directly on the argument advanced, particularly the more recent works that highlight the complex and situational nature of innovativeness and interpersonal communication (for example, King and Summers 1970; Summers 1971; Montgomery and Silk 1971).

Reprinted from D.F. Midgley, "A Simple Mathematical Theory of Innovative Behavior," *Journal of Consumer Research* 3 (June 1976):31-41, published by The Journal of Consumer Research, Inc.

The research approach evolved also from an interdisciplinary study—in its very broadest sense. Many writers have suggested that there are strong links between innovative behavior and the theories or models of mathematical epidemiology. These suggestions range from that of Ross (1915) to that of Vincent (1971) and include Coleman (1964), Bass (1969), and others. As well as showing the obvious fallacy of directly applying epidemic models to innovation, a review of the literature, reported at length in Midgley (1974), produced two important conclusions. The first of these was that the overall philosophy used by mathematical epidemiologists might be of considerable value. This is the specification of a set of differential equations on the basis of assumptions about "behavior." The second conclusion was that, in terms of advancing knowledge, a deterministic approach had more immediate value than a stochastic one. Innovative behavior is inherently more complex than that involved in epidemics, and there seemed little point in producing an abstract stochastic theory which could not be tested against reality.[1] Before the mathematics can be derived and discussed, however, it is first necessary to suggest some modifications to the theories of innovative behavior.

The Behavioral Theory

Many studies have documented the important role of interpersonal communication in influencing the adoption of new products since the pioneering work of Whyte (1954) and Katz and Lazarsfeld (1955), among them—Haines, 1966; Arndt, 1967; Sheth, 1971. Along with the earlier investigations arose the concept of opinion leaders, and implicit within this concept are two assumptions. First, that opinion leaders are an identifiable segment of the market and can therefore be influenced by the appropriate advertising. Second, that they do not exert influence for one product only, but in fact have generalized effect over a product category or several product categories. Recent studies suggest that overlap does exist, particularly between product categories involving similar consumer interests.

It has also been indicated that opinion leadership is widespread throughout the population (King and Summers, 1970) and further, that most of these "leaders" will have been influenced by others rather than by the mass media (Sheth, 1971). However, in the context of theory being developed here these generalized phenomena and concepts, while useful, are less relevant than the unique set of communications involved in the diffusion of any one new product.

While an indentifiable group of those most likely to exert personal influence may well exist, which members of this group actually discuss a *particular innovation* will depend on several factors. For any one of these individuals to be involved in the communication network it requires that

they: (i) have adopted before (ii) they meet a potential adopter and (iii) chance to discuss the product in question in the course of the conversation. All three of these are dependent on situational factors, thus rendering the communication network more random than previously stated. It would also appear that innovativeness is situationally dependent (see Summers, 1971).

In order to produce a more realistic description of the diffusion process it is proposed that we restate the concept of an opinion leader. Opinion leadership is in reality the combination of two traits, the first being *social involvement,* which is widespread and possessed to a greater or lesser degree by all members of society. That is, some people have more social contacts than others. The second trait is *product influence,* and is dependent on the degree of experience with the product in question. People who have adopted will have valid product influence which will be effective in conversations with potential purchasers. If these same people also possess a high degree of social involvement, then possibly they may influence a large number of other people to adopt. The necessary condition is adoption. An individual who has actually experienced the product will be able to exert greater personal influence than one who has not regardless of social involvement.

In essence, then, the view of opinion leadership, or product influence, put forward here is one of monomorphism. In this schema product influence is confined to one new product. If we were to study several innovations we might well discover a common group involved in communication, but if we wish to model the diffusion of one innovation then it is only relevant to study the influence messages relating to that particular product.

Lastly it can be hypothesized that *susceptibility to product influence* and *innovativeness* are also general traits possessed to a greater or lesser degree by all members of society.[2] However, the mathematical theory to be discussed here attempts primarily to model the operation of product influence within society, and in the context of a specific innovation. Only implicit provision is made for these other traits—by aggregating their effects into the model's parameters.

Since, as is often cited, most new products fail, we can infer that unfavorable interpersonal influence must be a common phenomenon. (See also Arndt, 1967; Robertson, 1971.) Hence, any useful theory of innovative behavior should allow for both the possibility of the spread of unfavorable influence, and the simultaneous communication of favorable and unfavorable information. It is plausible to state that once a person has adopted, and therefore experienced the product, he or she is able to form an opinion about it, which can be communicated to others. Depending upon how the individual perceives the product's performance this opinion may be for or against the innovation. For the sake of completeness we must also allow that some individuals may be neutral toward the innovation, either because they have no defined attitude toward it or because they do not converse with others.

It is necessary to hypothesize that product influence can be for or against the product in question, and that it can also take a neutral position. Remembering that those who have not adopted are assumed not to possess valid product influence, then it is possible to divide the population into four categories dependent solely upon the type of product influence they can communicate.

1. *Active Adopters:*—Those who have tried the product will give favorable information on it if the situation arises.
2. *Active Rejectors:*—Those who have tried the product, found it deficient, and will give unfavorable information on it if the situation arises.
3. *Passives:*—Those who have tried the product but do not give information or exert influence either way.
4. *Potential Adopters:*—All those yet to adopt.

To complete the new theory it is necessary to make a few comments on two other factors.

First, some individuals will adopt without having been influenced by others but by the direct effects of the media or free samples, for example. Once these individuals have adopted, though, they, too, can exert influence over potential adopters. Secondly, individuals may change their attitudes. A person may become bored with a topic, have a disagreeable experience with the product and change from being an active adopter to an active rejector, or, conversely, an active rejector may decide that he or she was mistaken and that the product is not as bad as was originally thought. All of these are changes within the individual and not caused by others, although they may subsequently affect others. It is just as reasonable to suggest that interpersonal influence may cause a subsequent change of mind. An active adopter may hear of the unfavorable experience of another person, re-evaluate his or her own experience, and switch to being passive or even an active rejector. Similarly this may occur for the other possible transitions of state. Adoption and the formation of a value judgment toward the product are not the end of the process; any individual may change from influence state to influence state as his or her attitudes change through personal experience or that of others.[3]

Figure 20–1 presents the theory of innovative behavior proposed above in the form of a system diagram. From this diagram it is a relatively easy task to derive a set of differential equations. It should be noted, however, that the transfers between adopter categories shown in figure 20–1 are only those allowed by the preceding behavioral assumptions. For instance, people may only *leave* the potential adopter state, and only by adopting the product. Furthermore, only the active adopters may persuade them to adopt. Potential adopters can, of course, adopt because of other factors not

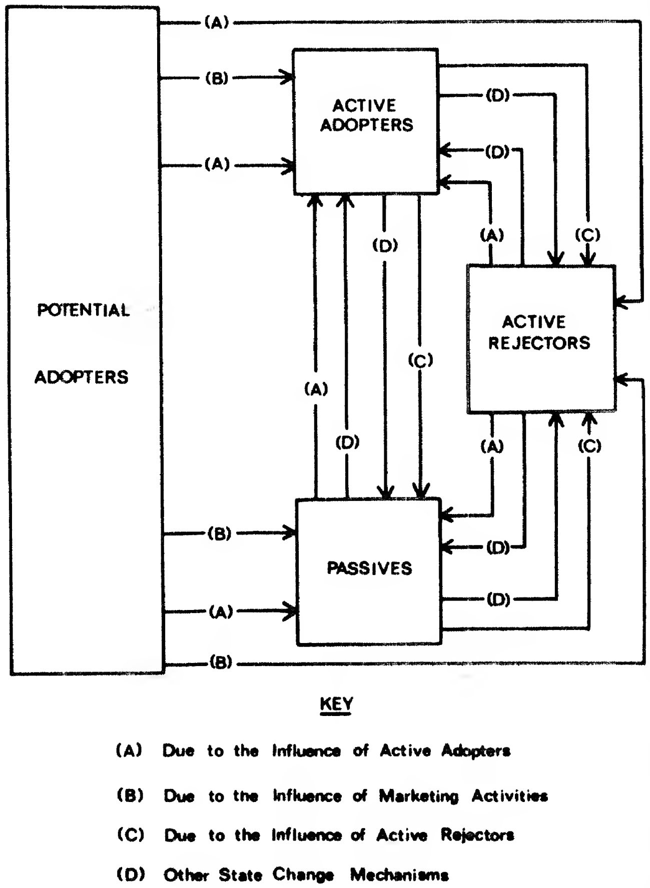


Figure 20–1. A Diagrammatic Representation of the Theoretical System.

related to interpersonal influence. Indeed it is necessary for a small proportion of them to do so in order to get the process started, but they cannot be persuaded to adopt by the active rejectors or the passives. Once a person has adopted the product then he or she can be persuaded to change his attitude toward it by conversations with either active adopters or active rejectors. The passives have no role in persuading others. They may only be influenced by others. An adopter may also change his or her attitude state because of internal effects such as boredom and the six possible transitions of this nature are allowed.

The result of this structuring of the influence mechanism is that the

effect of the active rejectors is indirect. If a large number of them exist they begin to persuade the active adopters that there may be something wrong with the product. If individuals encounter an increasing number of their social contacts telling them about the deficiencies of the product, they may well reevaluate their own position. Thus, a large number of rejectors subsequently produces a decrease in the number of active adopters and a slowing in the rate of adoption. The actual cumulative adoption curve may still rise because of the presence of the other, non-interpersonal influence mechanisms.

The active rejectors therefore have an inhibiting role while the active adopters are responsible for the phenomenon of "contagion." If a high percentage of the population is in favor of the innovation (having adopted it), then they are likely to persuade a proportion of the rest to adopt, thus increasing the percent in favor, and so on ad infinitum, provided that a similar proportion of these later adopters view the product as favorably. If not, then a different situation may well result. Most innovations will involve both favorable and unfavorable influence. According to the extent and dynamic behavior of each, a great variety of results will be produced. At this point it is difficult to extend the verbal description of the theory any further; to describe the variety of possible outcomes it is necessary to express the theory mathematically.

The Mathematical Theory

The new theory of innovative behavior results in a system which is less structured and in some ways more random than those proposed before. This allows the treatment of interpersonal communication by a simple mathematical expression of the form:

$$dY/dT = KXY, \text{ where } K \text{ is a constant.}$$

That is, the rate of transition of individuals in state X to state Y is proportional to the total possible number of pairwise meetings between these two groups (XY). If state Y contains active adopters and state X potential adopters, then indeed the rate of adoption will be proportional to the number of meetings between adopters and potential adopters, that is, proportional to the total (XY).[4]

The expression above can then be used for any transition between categories or states generated by interpersonal influence, and in terms of the system described in figure 20-1, these changes are not only between potential adopter and adopter but between all four categories. Having dealt with personal influence it is also necessary to derive an expression for other

mechanisms such as sampling, direct media-initiated purchasing, and boredom. These transitions do not involve the influence of others and therefore the XY term is inappropriate. In order to minimize mathematical difficulties the simplest possible formulation was used:

$$-dX/dT = JX, \text{ where } J \text{ is a constant.}$$

Put another way, a constant linear proportion leave one category and join another. Undoubtedly this is too simplistic an approach but it was thought desirable to reduce mathematical complexities. Further, the performance of this expression may provide valuable pointers to the best ways of improving the mathematical aspects of the theory. With the aid of the two basic expressions it becomes possible to formulate the system diagram in mathematical terms (figure 20-2) and then to write down the equations describing this system.

$$-dX/dT = XY(K_1 + K_2 + K_3) + X(J_1 + J_2 + J_3) \tag{20.1}$$

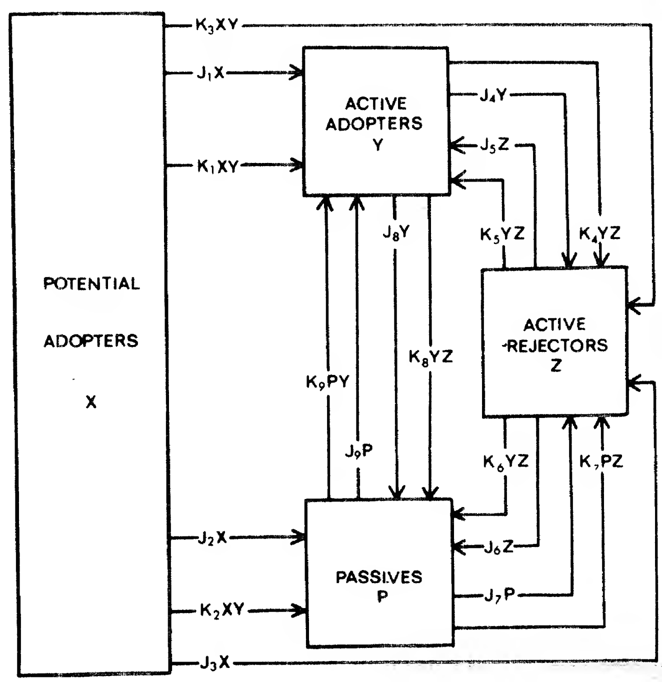


Figure 20-2. Mathematical Flow Diagram of the System

$$dY/dT = K_1XY + J_1X + K_5YZ + J_5Z + K_9PY + J_9P - K_4YZ - J_4Y - K_8YZ - J_8Y \quad (20.2)$$

$$dP/dT = K_2XY + J_2X + K_8YZ + J_8Y + K_6YZ + J_6Z - K_9PY - J_9P - J_7P - K_7PZ \quad (20.3)$$

$$dZ/dT = K_3XY + J_3X + K_4YZ + J_4Y + K_7PZ + J_7P - K_5YZ - J_5Z - K_6YZ - J_6Z \quad (20.4)$$

It will be noted that the number of potential adopters, X, can only decrease, and also that while the numbers in one or more of the other categories must therefore initially increase Y, Z and P may subsequently increase, remain stable or decrease.

Connecting these equations to actual data, (first purchase or market penetration), turns out to be quite simple. If C is the cumulative number of adopters, then

$$C = Y + Z + P \quad (20.5)$$

Also, since this is a closed system,

$$N(\text{population size}) = X + Y + Z + P \quad (20.6)$$

Equation 20.6 also demonstrates that as long as N can be regarded as a constant then for computational purposes one variable may be eliminated and equations 20.1 to 20.4 reduced to three equations in three unknowns.

All that remains are the initial conditions, and these are fairly easy to derive. At the moment a new product is launched ($T = O$) there are no adopters, only potential adopters. Therefore, Y, Z and P all equal zero, and the number of potential adopters equals the population size ($X = N$). This further means that the process must be initiated by the marketing mechanisms (J_1X, J_2X, J_3X), for which no adopters there can be no interpersonal influence. That is, the contagion terms K_1XY, K_2XY, K_3XY are all zero. This is in complete accordance with reality.

In summary, then, equations 20.1 to 20.6 together with the initial conditions form the mathematical theory of innovative behavior proposed here. The mathematics were specified from the system of behavior formulated earlier, a system which is thought to represent a more realistic description of the diffusion of new products. It remains now to discuss the performance of the theory against actual market situations. However, before this can be achieved it is necessary to outline briefly the methodology devised for solving the equations and fitting the solution to time series data.

The Methodology Used to Test the Theory against the Data

Equations 20.1 to 20.6 do not appear to have an analytical solution, that is, one which expresses the adoption curve as a function of the constants and time. However, this raises no difficulties because techniques for the numerical solution of differential equations are well established. In fact, most computer libraries have standard subroutines for such tasks. The subroutine used here was specially written for the purpose and was based on a Sarafyan variant of the Runge-Kutta method (see Lapidus and Seinfeld, 1971). The algorithm has been extensively validated against three standard integration programs and found to agree to a high level of accuracy. Further details are given in Midgley (1974).

Thus, given values for the constants, and performing the necessary computations, a time series can be produced for the adoption curve. However, since this is a completely new conceptualization we have no a priori knowledge of these constant values. Indeed we would expect them to vary widely from product to product. It is therefore necessary to determine these constant or parameter values by the familiar technique of finding the best fit between the theoretical solution and the data. As the equations cannot be solved analytically it is impossible to use the classical indirect optimization methods and, instead, a direct method must be employed.

The specific technique used was a modification of that originated by Hooke and Jeeves (1961) and described at length in Wilde and Beightler (1969). This method has been applied to the solution of a brand-switching model by Wormer and Weiss (1970). The objective function was the usual summation of the squared differences (errors) between the theoretical and actual time series: that is, between the time series for C and the actual cumulative adoption curves. Obviously it was required to minimize this.

Two problems with such a methodology are the amount of computer time required and interpreting the solutions produced. Developments subsequent to this particular research indicate that the former can probably be overcome, but the latter problem is likely to remain.

With an analytical technique such as linear regression we would be certain that the fit produced was the best obtainable, that is the "least square" fit. With direct pattern search techniques there are no such guarantees. The optimum solution may not be the best or "global" minimum, but a "local" minimum or a "saddle point." The pragmatic solution adopted for this research was to start from different sets of initial parameter values and perform several pattern searches, accepting the search which produces the best result. However, while the parameter values to be quoted later were established to a high accuracy these values relate to the "optimum" produced

and we cannot say that this is the best than can be achieved. There is a small but distinct possibility that a "better" or more global "optimum" might produce a totally different set of parameter values.[5]

Finally a third problem area exists, associated with the particular type of data used. The cumulative adoption curves were generated from consumer panel information and, as panels are only samples of the population, these curves are subject to statistical uncertainty. It would also be desirable if the parameter values obtained by the various optima could be stated with 95 percent or 99 percent error limits. In other words, if estimates of the population parameters could be made. It is possible that methods being developed at present may eventually help in this respect, but for the moment it is necessary to accept less perfect measures.

The Results

Consistency between Theory and Data

This was assessed by fitting the equations to the maximum amount of data available for each product. The results are given in table 20–1. The Kolmogorov-Smirnov test for the goodness of fit of empirical distributions measures whether the maximum discrepancy (D) between a theoretical and an actual cumulative frequency distribution can be reasonably thought to be a chance fluctuation from the theoretical distribution. (Massey, 1951).[6] In the worst of the six cases quoted we would expect a chance fluctuation from the theoretical distribution (the model) greater than or equal to that observed 20 times out of 100 (product 4). The values for the chi-square test are quoted in the same format.[7] A priori it was hypothesized that the residuals of the cumulated series would be positively autocorrelated, due to the smoothing and ordering effect of cumulation. The residuals of the first difference series were expected to be independent if and only if the theory was a reasonable description of the data.[8] The test used was that due to Durbin and Watson (1951).

Despite the usual reservations about the statistical measures employed it is evident that table 20–1 represents some support for the theory. The averaged squared errors are all low, the correlation coefficients all high and the Kolmogorov-Smirnov test would not lead to a rejection of the theoretical solutions obtained. Furthermore, five out of six solutions would not be rejected by the chi-square test and all of the residual series exhibit the hypothesized pattern of serial correlation.

Table 20–1
Testing the Theory against Panel Data: Values Obtained for the Statistics

Product Number	*Product Type*	*Number of Observations (Weeks)*	*Average Squared Error (%² per Week)*	*Correlation Coefficient*	*Kolmogorov-Smirnov Probability Level*[a]	*Chi-Square lative Level*[a]	*Results of the Durbin-Watson Test on the Residuals (0.05 Level)*: *Cumulative Series*	*First-Difference Series*	*Total Panel Size*
1	Toothpaste	42	0.188	0.998	89.3	99.8	Positive[b]	Independent	368
2	Confectionary	20	0.049	0.998	99.9	5.8	Positive	Independent	545
3	Detergent	26	1.026	0.998	53.3	44.3	Positive	Independent	490
4	Detergent	23	3,286	0.994	20.0	0.0	Positive	Independent	438
5	Toothpaste	27	0.131	0.996	49.8	85.8	Positive	Independent	652
6	Biscuit	35	0.465	0.999	40.0	9.8	Positive	Independent	835

[a]Read as the probability of occurrence, under the theoretical hypothesis, of a value greater than or equal to that observed. The probability levels have been rounded to one decimal place.

[b]That is, positive serial correlation.

Predictive Ability

Assessments of the consistency between theory and data are a relatively weak test of the theory's value. To arrive at a more objective evaluation it is necessary to pose the question of whether the theory is capable of predicting human behavior in this area. However, with eighteen parameters it is necessary to have at least twenty observations, that is one degree of freedom. Unfortunately it was impossible to obtain longer data series than those detailed in table 20–1, and hence only two meaningful forecasting exercises could be undertaken with any confidence. These were performed by fitting the theory to twenty weeks' data for Products 1 and 6 and then projecting the results ahead for the remaining available data.

For Product 6 the resulting forecast was excellent, very little different from the data curve for the remaining fifteen weeks. However, for Product 1 the prediction consistently overestimated the actual adoption curve by significant amounts. A plausible explanation of this failure is that the panel concerned contained the smallest number of actual purchasers encountered over the various estimation periods used.[9] A combination of random fluctuations in such small numbers may very well have biased the predictions in an upward direction. In the context of minimum chi-square estimates of the parameters of stochastic brand choice models Blattberg and Sen (1973) conclude that samples of more than 500 respondents are needed. These authors also suggest that samples of less than 500 bias the parameter estimates upward. As far as the development of new product models is concerned the results discussed here suggest that a much larger panel size is necessary, particularly where reliable predictions are needed on the basis of only a few weeks' observations.

Results comparable to those presented above could equally well have been generated by a two or three parameter growth model (Gompertz logistic). The statistics have been quoted merely to show that this theory yields a reasonable level of performance against empirical data. Where the conceptualization does differ from simple models is that it is based on extant knowledge and therefore provides more useful insights into consumer behavior. One such insight can be demonstrated by looking at the results in terms of the dynamics of the adopter categories rather than the cumulative adoption curve.

Adopter Dynamics

The numbers of people in the various categories at any point in time can easily be graphed, and figures 20–3 and 20–4 illustrate two typical products in this manner. The sum of the three adopter categories is, of course, the

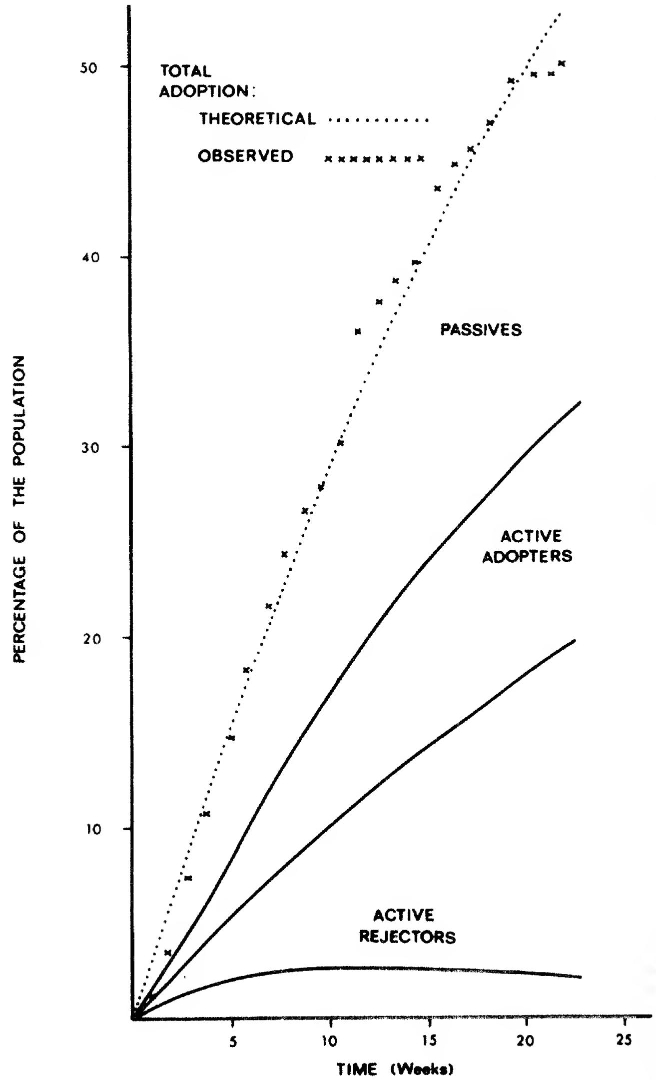


Figure 20-3. Adopter Category Behavior for Product 4

cumulative adoption curve and this is represented in the figures by a dotted line. The potential adopter category has been omitted for the sake of clarity; as the results are expressed in percentages of the population, at any point in time the numbers in this category are 100 minus the number of adopters. The graph of those still to adopt is therefore the mirror image of the cumu-

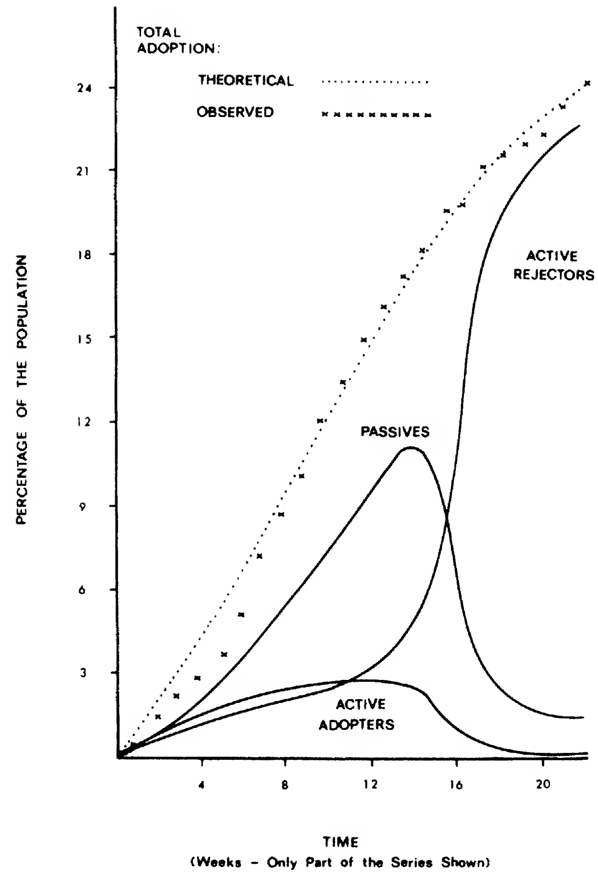


Figure 20-4. Adopter Category Behavior for Product 6

lative adoption curve. Inspection of figure 20-3 leads to the conclusion that the product in question was a successful innovation in the sense that its introduction produced very few who disliked it, these active rejectors remaining at around two percent of the population. On the other hand, Product 6 (figure 20-4) was unsuccessful in that by the twentieth week after launch most of the population viewed it unfavorably.[10] However, these conclusions are inferential as there are no quantitative data on the behavior of the categories available, only on the behavior of the total adoption curve.

To prove or disprove these inferences conclusively would require the development of new methodologies, as the adopter categories relate to situational states of mind. A survey of the population might well measure the general propensity of individuals to fall into one or another influence state,

but which state they actually fall into for a particular innovation is uniquely determined by that innovation, and the influence messages associated with it. One viable approach may be to extend the panel questionnaire to monitor opinions on the product; another, to hypothesize that the active adopters will repeat their purchases while the active rejectors will not. This would enable a prediction of the curve for total sales to be made as well as that for first purchases.

It is also true that results are variable, in the sense that minor differences in the shape of the adoption curve for different products apparently produce major differences in the curves for the individual adopter categories.[11] It is possible to argue that the phenomenon the theory attempts to model is by nature itself unstable, involving as it does the process of social contagion, but more importantly it is possible within the scope of this article to present some evidence which relates the category of behavior to reality.

This evidence is subjective and much less conclusive than might be desired, but there are problems in collecting even these data. The most informative cases relate to failure, and management is not overly willing to discuss such events. However, Product 6 (figure 20-4) provides one striking illustration. There it will be seen that the launch was fairly successful up to and around week 12. At that point the numbers in favor of the product begin to decline. This decline is followed by a similar fall in the number of the passives at around week 15. On the other hand from week 12/13 onward the active rejectors, that is, those against the product, begin to grow rapidly until by week 35 they account for almost all the first purchasers. The explanation of this phenomenon is that at somewhere around the point detailed by the solution it is known that the real-world distribution system encountered problems. First purchasers of the product became unable to repeat their purchase of it and this in all probability resulted in a swing of opinion against the product, a swing which the theoretical solution appears to depict. There is similar evidence linking the category solutions of two other products to management's perception of what actually happened.

Parameter Estimates

The fourth and last set of results are the parameter values obtained in fitting the theory to data (the values of the K and J constants). Comments have already been made about the difficulties of interpreting these values. They are presented here primarily so that other workers may replicate the solutions already given (table 20-2). To that end the data are also presented (table 20-3). Tentatively at least, it would appear that there are regularities in the parameter values given. For instance, in five out of six cases J_1 is greater than J_3. That is, marketing activities such as advertising and sam-

Table 20–2
Values Obtained for the Constants

Constant	*Product 1*	*Product 2*	*Product 3*	*Product 4*	*Product 5*	*Product 6*
K_1	0.0002649	0.0026657	0.0000046203	0.0	0.0031586	0.0011522
K_2	0.00014715	0.00017043	0.0037724	0.0	0.0	0.00053173
K_3	0.0	0.25922	0.0033824	0.0	0.000063772	0.00058022
K_4	0.0	0.15977	0.24998	0.17	1.0708	0.9509
K_5	1.8675	0.53491	0.58996	0.85	0.51314	0.99937
K_6	0.21877	0.0	0.71995	0.79	0.0040077	0.99937
K_7	0.60487	0.29157	0.1156	0.62	0.5385	0.1318
K_8	0.0023652	0.83766	0.55736	0.33	0.86016	0.95055
K_9	0.0	0.58957	0.059996	0.0	0.84789	0.040474
J_1	0.0031287	0.0	0.012717	0.014231	0.0027672	0.0063958
J_2	0.00035256	0.0043279	0.002858	0.0098	0.0	0.0012554
J_3	0.00058239	0.0	0.0	0.0096	0.0000058759	0.0019804
J_4	0.76569	0.96179	0.00004668	0.0	0.0049829	0.45731
J_5	0.0	0.36937	0.0	0.0	0.59054	0.038308
J_6	0.060907	0.59556	0.36591	0.79	0.19944	0.034774
J_7	0.21473	0.0012502	0.47997	0.53	0.00313376	0.36977
J_8	0.63668	1.0413	0.89987	0.95	0.38749	0.17982
J_9	0.19332	0.0014060	0.12999	0.17	0.27863	0.71893

Table 20-3
The Data: Cumulative Adoption
(percentage of households)

Week	*Product 1*	*Product 2*	*Product 3*	*Product 4*	*Product 5*	*Product 6*
1	0.20	0.81	2.53	1.03	0.46	0.36
2	0.20	1.14	4.69	3.10	1.07	1.44
3	0.40	1.30	7.20	7.23	1.38	2.16
4	1.00	1.46	9.23	10.80	2.30	2.99
5	2.00	1.68	11.44	14.69	2.76	3.83
6	2.59	2.68	14.58	18.05	3.83	5.27
7	3.19	3.02	19.63	21.59	5.21	7.31
8	3.59	3.69	22.24	24.51	5.52	8.86
9	3.79	4.08	23.55	26.82	6.13	10.42
10	4.39	4.58	24.86	27.89	6.60	12.46
11	4.59	5.09	26.78	30.02	7.06	13.65
12	4.79	5.94	29.09	36.27	8.13	15.33
13	5.19	6.77	31.98	37.50	8.44	16.53
14	5.79	7.49	33.91	38.73	9.36	17.49
15	6.19	7.84	36.67	39.96	9.66	18.56
16	7.29	8.02	37.88	44.10	9.97	20.00
17	7.71	8.81	38.68	45.21	10.12	20.24
18	7.92	9.54	39.28	45.88	10.28	21.32
19	8.54	10.46	40.61	47.22	10.74	22.04
20	9.01	10.83	41.63	49.32	10.74	22.40
21	9.87		42.04	49.54	11.20	22.99
22	10.73		42.45	49.54	11.50	23.95
23	11.16		43.06	50.23	11.66	24.91
24	11.50		44.08		11.96	25.39
25	11.73		44.29		12.42	26.11
26	12.17		45.92		13.19	26.95
27	12.83				13.96	27.54
28	15.25					28.26
29	15.25					29.58
30	15.50					30.30
31	15.75					31.38
32	16.62					32.46
33	17.41					32.57
34	17.68					33.05
35	18.21					34.13
36	18.77					
37	19.57					
38	20.11					
39	20.38					
40	21.20					
41	21.74					
42	22.01					

pling create more people in favor of the product than against it. The combined effect of J_1 and J_2 (the effectiveness of marketing activities in creating neutral purchasers) is always greater than J_3. All of which is extremely plausible, and indeed heartening for marketing management. In making further comparisons another note of caution must be added. It is not pos-

sible to compare directly the value of a K constant with a J constant as the former is multiplied by the product of two variables while the latter is only multiplied by a single variable. To make these comparisons, a detailed analysis of the transfers between categories is needed (that is, an analysis of the individual terms in equations 20.1 to 20.4). Only the broad conclusions of this analysis are given here.

It would appear that the overall importance of interpersonal influence in creating first purchasers ($K_1XY + K_2XY + K_3XY$), vis-a-vis that of marketing activities ($J_1X + J_2X + J_3X$), is highly product-dependent. In two cases interpersonal communication was more effective than the media; in one case, the effects were equal; and, in the remaining three cases, marketing activities were more influential. However, a subtler analysis, based on the effectiveness of communications in creating people in favor of the product, revealed a more complex situation. In two cases word-of-mouth was more effective in creating the favorable, or active, adopters, and in one case the media were much more influential. For one product both interpersonal communication and marketing activities were equally effective in generating favorable opinions, while in the remaining two cases marketing activities were producing those in favor of the product and word-of-mouth was producing those who disliked it. Nor is this quite the complete picture as the comparisons are based on peak values and the dynamics of the situation have an important bearing on the end result.[12]

For those who *had* purchased, interpersonal influence, and specifically unfavorable influence, appeared more effective in changing attitudes. This would suggest, albeit speculatively, that once a person had adopted, then other other peoples' opinions are far more likely to change his/her attitudes than are internal factors such as his/her perception of the product. However, care in interpretaton is needed here because the opinion of others is likely to have been formed by internal factors (especially a disagreeable experience with the product). In essence, a small proportion of the population experiencing poor product performance can alter the attitudes of the majority and eventually slow the rate of adoption. It is therefore dangerous to assume that because an effect is small, or of lesser importance than another, it can safely be removed from the model. It may well have been the sole cause of the final result. Whatever these comments do or do not demonstrate they do show that it is no longer possible to emphasize one type of communication at the expense of another. Any theory of innovative behavior must incorporate and relate all channels of communication and influence.

General Comments

In conclusion, some more general remarks on the practical aspects of the approach taken, and more importantly on the nature and reliability of the

estimates so obtained, are appropriate. First, as indicated before, direct search optimization is a relatively costly and time-consuming exercise. It is difficult to give precise figures as the computer time necessary is problem dependent, ranging from 18 to 163 minutes (*CDC* 7600, each solution of the equations taking an average of 0.5 seconds), although the majority of runs were nearer the lower figure than the upper. As previously stated, other exercises from different starting values were also necessary. Now it must be stressed that the research was of an exploratory nature, and as such encountered numerous problems, and there is little doubt that with the experience gained the times quoted could be substantially, even drastically, reduced. A change to a more recent technique would probably produce greater benefits, either a specialized least-squares algorithm such as that due to Peckham (1970) or one of the many variants of the more general Davidon Fletcher-Powell method (see Wilde and Beightler 1967). Seemingly, a "mixed" approach may be best—using direct search techniques to produce initial improvements, then the Davidon Fletcher-Powell method to "climb down" to the region of the minimum, and finally the fitting of a response surface.

Secondly, and perhaps more important given the aim of the research, the results obtained and estimates produced appear reliable. While it is possible that the respective optima could be further improved by processing to higher accuracies, this would have no material effect on the resultant time series, and it is therefore difficult to justify the computer time necessary.[13] The original estimates given in the objective function of 10^{-4}. That is, the program terminated when unable to find improvements greater than this figure. The estimates given here are established to 10^{-5} (or higher), and this factor is responsible for the differences between the two sets of results. These differences are in the majority of cases very slight, and in particular the time series for the adopter categories remain substantively the same as for the earlier results.[14]

Both the earlier and the present results have been subjected to various experiments which revealed, as is often the case in optimization, that the contours of four of the objective functions were asymmetric and that the parameters of all the six functions showed some degree of interaction or correlation. Also while the objective functions are reasonably convex with respect to the parameters K_1 to K_3 and J_1 to J_3, for the other parameters the functions are shallow. All of which would be discouraging were it not for the following.

It appears that the behavior of the time series for the various adopter categories is extremely robust to alterations in the parameter values. This has been demonstrated by a special program which ascertains the magnitude of a shift in the parameter values (singly, as in a univariate search, and from the values at the optimum) which results in the time series for the three adopter categories being on average between 5 and 10 percent different from the "optimum" time series. By 10 percent is meant 10 percent of each

figure, not 10 percent absolute adoption, and this effectively means only a moderate distortion of the behavior of these time series.

In the vast majority of cases these alterations in the parameter values were found to be substantial and coupled with significant increases in the objective function values.[15] Since such variations would have been detectable to the optimization technique it is possible to conclude that the patterns of "success" or "failure" depicted by the results, and for example in figures 20-3 and 20-4, are stable throughout the region of the respective optima. Put another way, for the time series to be radically distorted requires a significant increase in the objective function value and an unlikely change in the parameter values. The only problem remaining is that within these limits the parameters, or at least some of them, can be altered with slight effect on the objective and without any effect on the time series. This is why some further processing was thought necessary, in essence to "pin down" the estimates more closely. It is now considered that the quoted parameter values are accurate to approximately (±) 8 percent Product 1, 4 percent Product 2, 0.05 percent Product 3, 0.01 percent Product 4, 8 percent Product 5, 0.6 percent Product 6.[16] As a full discussion of the above topics is beyond the scope of this paper, all the results to date are to be made available to interested parties as a working document (see Midgley 1976). To summarize, the behavior of the adopter categories is unique to each product and its associated optima, and furthermore is stable against quite severe changes in the parameter values. The results cited can therefore be viewed with some considerable degree of confidence.

Conclusions

It has been shown that it is possible to express the behavioral theory relating to innovative behavior as a set of differential equations. The solutions obtained in applying this theory to consumer panel data suggest that the theory has some value. It would also appear that there are good grounds for inferring that the system solutions do relate to actual events. The chief reason for these measurements being as yet difficult and costly would appear to be due to the nature of the optimization technique used, rather than any major defect in the theory. It has also been concluded that it is possible to overcome these problems, facilitating more accurate solutions and the processing of data for a wider range of products. Such improved techniques would also allow the generation of estimates of population parameters, with associated errors, and hence enable interproduct comparisons of these parameter values to be made. This may be the single most worthwhile contribution the theory could make in the future. The present results indicate that the solutions lie close to the respective global optima and that there are

regularities in the parameter values obtained for these solutions. Once it becomes possible to assign confidence limits of some kinds to these values, then it will be feasible to analyze them on a much more systematic and rigorous basis and hopefully to discover previously unknown facets of human behvior.

Further it may be practical to collect data on the individual adopter categories proposed here and hence test the theory more extensively than can be achieved with a single adoption curve. Indeed, if it proves possible to monitor these categories accurately, then the model itself could be fitted simultaneously to the three curves. This should provide more precise measurements of behavior. Alternatively, a less costly approach might be to incorporate a repeat purchase expression, hence modeling sales and adoption.

Finally, an exciting possibility is that of taking an optimal control theoretic approach to the problem of parameter estimation, except that in this case the parameters would no longer represent constants but would become variables themselves. By applying such methods (see Boltyanskii, 1971), an optimal "control history" or time series for each parameter would be generated, and could conceivably be linked to the advertising history for the product, and so on. While the potential in this area appears immense it must be admitted that the computational complexities are also severe.

In all probability the mathematical theory proposed is at best only a first approximation to the actual processes of innovative behavior. What the results given here do demonstrate is that it is possible to quantify complex areas of human behavior, and in so doing generate both new knowledge and feasible lines of future research. The fact that this methodology has partially succeeded indicates that it will shortly become possible to derive more realistic and wide ranging mathematical theories for this aspect of consumer behavior.

Notes

1. Within mathematical epidemiology it has proved extremely difficult to specify the differential equations in a stochastic form capable of yielding an analytical solution. Further, as in other fields, the alternative simulation approach has encountered the problem of relating the results to real data. On the whole, epidemic theories and models are simpler than that proposed here, primarily because the phenomenon itself is simpler, and hence the comments above. However, it should be noted that these are addressed to models which explicitly recognize that the process is dependent on the outcome of meetings between members of different groups in the

population: that is, they are dependent on contagion or, in the case to be discussed here, interpersonal influence. The comments do not apply to stochastic models in general.

2. The former by analogy to the trait "susceptibility to social influence" hypothesized in the field of social psychology. (See McGuire, 1968 and 1969.)

3. By influence states are meant the various adopter categories mentioned previously.

4. For the sake of simplicity homogeneous mixing of the population is assumed.

5. However, by definition it would also produce a better fit between theory and data.

6. The reasons for using the cumulative form of the adoption curve lie mainly in the nature of consumer panel data.

7. The process of interpersonal influence implies that the number of adopters in one time period is dependent on the number in the preceding period. Phillips (1971) suggests that in such cases the chi-square test is overly rigorous.

8. The first difference series represent the observed and expected number of adopters in each week. The residuals of the two series can therefore be thought of as the observed error between theory and "raw" panel data.

9. Despite the panel having 490 members, only 27 had adopted by week 16 as opposed to 167 for Product 6.

10. Of the 6 products two were "successful" (1 and 4), and the remainder "unsuccessful."

11. On the other hand, there is a significant difference between the gradients of the two adoption curves presented here, and this may be a part explanation of the differences in the solutions. It should be noted that figures 20-3 and 20-4 are not to the same scale.

12. The analysis of channel effectiveness is discussed in more detail in Midgley (1975).

13. Which is not meant to imply, of course, that the possibility of locating a new and completely different minimum does not still exist.

14. Indeed for three cases these series remain unchanged.

15. For instance in the case of Product 6 the smallest change was in K_5 and was 0.4%, this produced a 38% increase in the objective function value, and an average 8% "distortion" in the time series values. The smallest change in the objective was 6% and associated with a 43% alteration in K_3, and a 7% "distortion."

16. The parameter values are quoted to five figures solely that others may replicate the solutions obtained to date, that is, may generate identical

optima. To an approximate error in the objective function of 2×10^{-5}, 5×10^{-7}, 4×10^{-8}, 5×10^{-7}, 5×10^{-6}, and 4×10^{-5}, respectively.

References

Arndt, J. *Word-of-Mouth Advertising: A Review of the Literature.* New York: Advertising Research Foundation Monograph, 1967.

Bass, F.M. "A New Product Growth Model for Consumer Durables," *Management Science,* 15 (January 1969), 215–227.

Blattberg, R.C., and S.K. Sen. "An Evaluation of the Application of Minimum Chi-Square Procedures to Stochastic Models of Brand Choice," *Journal of Marketing Research,* 10 (November 1973), 421–427.

Boltyanskii, V.G. *Mathematical Methods of Optimal Control.* New York: Holt, Rinehart / Winston, 1971.

Coleman, J.S. *Introduction to Mathematical Sociology.* Glencoe: The Free Press, 1964.

Durbin, J., and G.S. Watson. "Testing for Serial Correlation in Least Squares Regression: II," *Biometrika,* 38 (June 1951), 159–178.

Haines, G.H. "A Study of Why People Purchase New Products," in R.M. Haas, ed., *Science, Technology and Marketing,* Chicago: American Marketing Association Conference Proceedings, Fall 1966.

Hooke, R., and T.A. Jeeves. "Direct Search Solution of Numerical and Statistical Problems," *Journal of the Association for Computing Machinery,* 18 (April 1961), 212–228.

Katz, E., and P.F. Lazarsfeld. *Personal Influence.* Glencoe: Free Press, 1955.

King, C.W., and J.O. Summers. "Overlap of Opinion Leadership across Consumer Product Categories," *Journal of Marketing Research,* 7 (February 1970), 43–50.

Lapidus, L., and J.H. Seinfeld. *Numerical Solution of Ordinary Differential Equations.* New York: Academic Press, 1971.

Massey, F.J. "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association,* 46 (March 1951), 68–78.

McGuire, W.J. "Personality and Susceptibility to Social Influence," in E.F. Borgatta and W.E. Lambert, eds., *Handbook of Personality Theory and Research,* New York: Rand McNally, 1968.

———. "The Nature of Attitudes and Attitude Change," in G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology Vol. III,* (2nd Edition) Reading, Mass.: Addison-Wesley, 1968.

Midgley, D.F. "A Quantitative Theory of Innovative Behaviour," unpublished Ph.D. thesis, University of Bradford Management Centre, 1974.

——— . *Managing New Products.* Cranfield: Cranfield Research Papers in Marketing 2, 1975.

——— . *Further Results in the Application of a Theory of New Product Adoption,* Working Paper, The School of Marketing, The University of New South Wales, 1976.

Montgomery, D.B., and A.J. Silk. "Clusters of Consumer Interests and Opinion Leaders' Spheres of Influence," *Journal of Marketing Research,* 8 (August 1971), 317-321.

Peckham, G. "A New Method of Minimizing a Sum of Squares without Calculating Gradients," *Computer Journal,* 13 (1970), 418-420.

Phillips, L.D. *Bayesian Statistics for Social Scientists.* London: Nelson, 1973.

Robertson, T.S. *Innovative Behavior and Communication.* New York: Holt, Rinehart / Winston, 1971.

Rogers, E.M. *Diffusion of Innovation.* Glencoe: The Free Press, 1962.

Rogers, E.M., and F.F. Shoemaker. *Communication of Innovations,* Glencoe: The Free Press, 1971.

Ross, Sir, R. "Some a Priori Pathometric Equations," *British Medical Journal,* March 27th, 1915.

Sheth, J.N. "Word of Mouth in Low-Risk Innovations," *Journal of Advertising Research,* 11 (June 1971), 15-18.

Summers, J.O. "Generalized Change Agents and Innovativeness," *Journal of Marketing Research,* 8 (August 1971), 313-316.

Vincent, M. "The Recognition and Systematic Study of Phenomenon of Diffusion and Contagion," Proceedings, Congress of the European Society for Opinion and Marketing Research, 1971.

White, W.H. "The Web of Word-of-Mouth," *Fortune* (November 1954).

Wilde, D.J., and C.S. Beightler. *Foundations of Optimization,* Englewood Cliffs: Prentice-Hall, 1967.

Wormer, T.A. Van, and D.A. Weiss. "Fitting Parameters to Complex Models by Direct Search," *Journal of Marketing Research,* 7 (November 1970), 503-512.

21 New-Product Sales Forecasting Using the Hendry System

Vithala R. Rao

When historical data on switching among brands in a product category are available, a variety of techniques can be employed for forecasting sales of established as well as new brands. These techniques are based on some conceptualizations of brand-switching behavior at the level of market as a whole. While several alternative conceptualizations are published in the literature, this chapter describes the system of models developed by the Hendry Corporation and shows its application to sales forecasting for a new brand.

At the outset, it must be noted that the Hendry system is proprietary; therefore, published information is necessarily not complete. It has been the focus of much academic debate amid several published testimonials from industry users. The discussion of this chapter is based on an understanding of the Hendry system developed over the years.[1]

This chapter is organized into four sections. The first presents an overview of the Hendry system. This is followed by a discussion of the data requirements for implementing the system. The third section shows how it may be used for forecasting sales of a new brand. The method is illustrated with an actual analysis. The final section points out some advantages of this approach vis-a-vis other stochastic model-building efforts. The technical details are presented in appendix 21A. The chapter is supplemented by another appendix prepared specifically for this chapter by the Hendry Corporation. Appendix 21B lists the required data input and corresponding output for a specific example.

The Hendry System

The Hendry Corporation has developed a system of models for describing consumer purchase patterns under various conditions. The system applies to frequently bought consumer products. Simple and parsimonious input data are used to estimate the parameters of the system. These estimates may then be used to predict sales of new brands in a product category under prespecified marketing-mix scenarios.

The literature on these models is very sparse. In fact, the only materials published that refer directly to the Hendry methodology are Hendry Corporation (1970, 1971) and a restatement by Kalwani and Morrison (1977) which explains some of the important concepts underlying the Hendry methodology.[2]

Major Concepts

At the core of the system is a stochastic model of brand choice. This model incorporates the strength of consumers' brand preferences as reflected in market shares and repeat-buying rates. Two important notions are introduced. The first is that of equilibrium, defined by stable market shares and consumer preferences. The second is the notion of a "directly competing set," that is, those items (brands) that are perceived by consumers as potentially filling the same "need."

Laws of the Hendry System

If one assumes equilbrium and that the set of directly competing items has been correctly identified, two "laws" of switching are expected to hold. The first (the law of detailed balancing) says that the aggregate switching from item i to item j is equal to that switching from item j to item i. The second (the law of proportional switching) says that the switching between any two items is proportional to the product of their market shares. The constant of proportionality is the same for all pairs of items in the set of directly competing items and, in fact, is an index of the level of switching (termed the *switching constant* and denoted by K_w).

Market Partitioning

The switching relationships are used to identify a hierarchial market structure, called *market partitioning,*[3] of any given product category. This idea provides for the possibility that not all available items in a category may compete with each other to the same extent. It implicitly assumes that substitutability (as measured by switching behavior) varies across certain subsets of items which may have common features along one or more dimensions. Thus, the aim of market partitioning is to hierarchically divide the total set of all items in a product category such that switching levels are different at different levels of the hierarchy and the observed switching is matched with theoretically described switching. The theory of entropy is used in postulating the theoretical model.

Although neither specified nor required in the Hendry system, a correspondence between the levels of hierarchy and degree of switching could be intuitively appealing. Thus, for example, in figure 21-1 the highest level of switching may be expected to occur among the different flavors of a given type/brand. At the next level, different brands of a given type form a set with possibly a lower level of switching among them. At the next level, different types form a set with an even lower level of switching. Knowledge of the appropriate market structure is critical for forecasting the share or sales of a brand.[4]

The application of the Hendry partitioning approach to several product categories revealed one of two forms of partitioning structures: nested and mixed-mode (Rubinson, Vanhonacker, and Bass 1980). In the nested structures, the buyer is deemed to organize the alternatives in the market in a sequential manner. Either the type of product or the brand name is considered to be the primary level of partition. These structures are respectively called type-primary and brand-primary market partitions. The illustration in figure 21-1 is a type-primary partition.

The mixed mode structures when two product characteristics, say brand name and type of product, simultaneously form the primary partitioning level. An example of such a structure is shown in figure 21-2. In this structure, consumers perceive the total market (product category) as homogeneous in terms of not only brands but also product types. Apparently, some markets do not empirically conform to nested partitioning structures; this

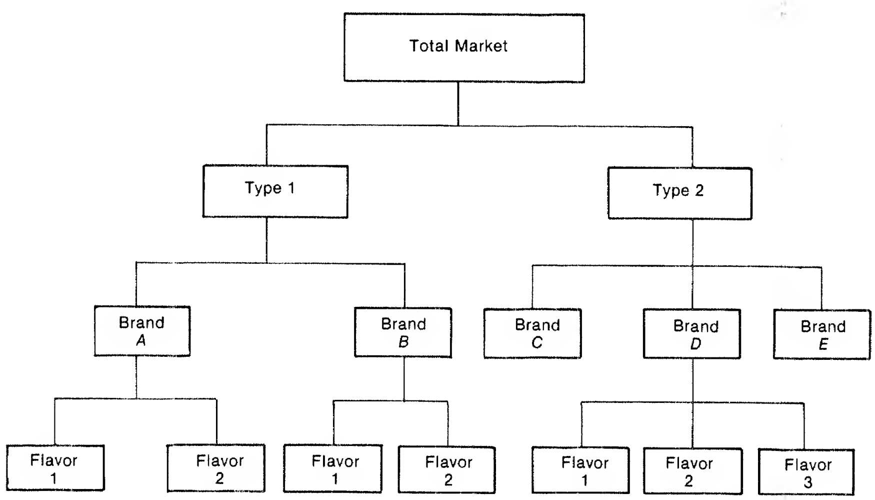


Figure 21-1. A Hypothetical Market-Partitioning Structure

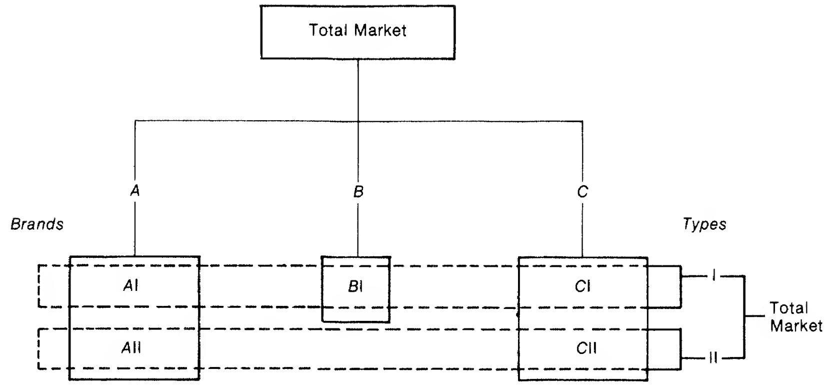


Figure 21-2. A Hypothetical Mixed-Mode Partitioning Structure

fact necessitated the Hendry people to introduce the mixed-mode partitioning construct. See Vanhonacker (1980) for an application of Hendry partitioning methods.

Identification of appropriate structures is critical in the Hendry analysis. Their method is briefly discussed in appendix 21A. The reader may note that there is no hierarchy of switching constants in the empirical determination of market partitions according to the Hendry theory. The entropy theory is applied to describe the switching behavior observed in the market. A more complete discussion can be found in Kalwani (1979).

Several alternative approaches to market-partition determination are emerging in the literature.[5] These include the efforts by Rao and Sabavala (1978, 1979); Day, Shocker, and Srivatsava (1979); Shocker and Day (1976); Urban (1977); and Urban, Johnson, and Brudwick (1979). In particular, the Rao-Sabavala method utilizes cluster analysis of similarity measures based on the switching matrix to obtain a hierarchy and maximum-likelihood methods to determine the parameters of the hierarchy.

Par Share

The Hendry system is particularly useful in estimating the market share of a new brand prior to introduction in the marketplace. This estimation procedure uses the concept of *par share*. "Par share" refers to the mature performance (that is, market share) expected for a new brand perceived to be of equal value to consumers as the existing brands in the marketplace. The concept of equal value is described as the *par* brand and is defined by the profile of marketing decision variables (for example, price, advertising, and

distribution level) of the par brand. Operationally, the profile is measured by the weighted averages of prices, advertising, and distribution of the existing brands in a given market partition in which the new brand is expected to compete. The weights used are the respective market shares of the existing brands in the partition.

While not all the details of the relationships are revealed, the Hendry people developed a mathematical model to predict the par share. This is known as the PARSHARE model. Presumably, it contains a set of relationships (empirical or hypothetical) between the marketing-mix decision variables and market share in the context of a market partition. The reader might note that the market-partition context of the new brand is closely tied to the entropy theory described in the technical details published by the Hendry Corporation. The PARSHARE model is quite versatile; it can be (and, in fact, is) used to predict the mature performance of brands with profiles other than the par brand. Further, it has the capability of determining the "optimal" marketing decision variables for a new brand in a partition, optimal with respect to meeting a set of predefined objectives for the brand (for example, market share, profit level, and so forth).

While we have described here the model for mature performance of the new brand, other aspects of the Hendry system deal with the dynamics (transition) of the new brand.

Data Requirements for the Hendry System

The preceding discussion dealt with the PARSHARE model, only one of the several subsystems of the Hendry system of models. Their system also includes tools/models for analyzing decisions relating to the marketing-mix variables (for example, advertising, price, price promotion, distribution level, product sampling, and so on) and also for studying the expected dynamics of the marketplace resulting from implementation of these decisions for brands.

The critical component of the Hendry system appears to be the determination of the appropriate market partition for a product category. The market partition is described by a number of switching constants (K_w's), which are determined by imposing the theoretical entropy model. The procedure is described in appendix 21A. The data required for this portion of the Hendry system are a transition matrix of brands (or items) in a product category and some hypotheses (conjectures) of possible market partitions. The transition data can be obtained either from a consumer panel or from a survey with questions such as: What brand do you use now? What brand did you buy last time? What brand did you buy the time before last? What brand did you buy time before that? The description of the brand is

obtained in its entirety, that is, pack size, flavor, place bought, and so on. Presumably, a stable time period is chosen for determining the market partitions. The derived market partitions (that is, market structure) are assumed (and also sometimes claimed) to be stable for a product category for a fairly long time.

The data necessary for the other parts of the system include prices, advertising-expenditure levels, price promotions, distribution levels achieved, and so on for all the items defined in the market structure. Some assessment of the quality (or effectiveness) of advertising of various brands is also inputted into the models. In order to convert the market shares to other outcome measures, the system also requires data on gross margins for brands and external estimates of total demand for the product category. Naturally, these data have to time-dependent. It is a conjecture that fairly long historical data on these variables are required in order to estimate stable response functions for use in various components of the models.

Forecasting the New Brand's Sales

The Hendry methodology is useful for forecasting sales of brands in a competitive setting where the primary demand of the product can be predicted reasonably well and where the major problem is predicting the share of market that a specific brand can capture. The forecasted market shares refer to the equilibrium state of the market. The methodology can provide answers to three questions of interest:

1. Given the market structure, what share can a new item expect to achieve under the assumptions that it is on "par" with respect to quality and marketing effort? (This share is referred to as the par share.)
2. In the above situation, what share can the new-product entry achieve under alternative marketing strategies? In fact, can one find "optimal" (with respect to share of market or financial return) levels for the price, advertising, and promotional efforts?
3. What share can be expected for an existing brand under a change in marketing strategy? Again, what would the "optimal" levels of marketing support be?

To answer these questions, we need two of the components in the system described conceptually above. The first is the mechanism by which a new items draws from existing items in the set. The second is a set of response functions relating purchase behavior to the marketing decision variables.

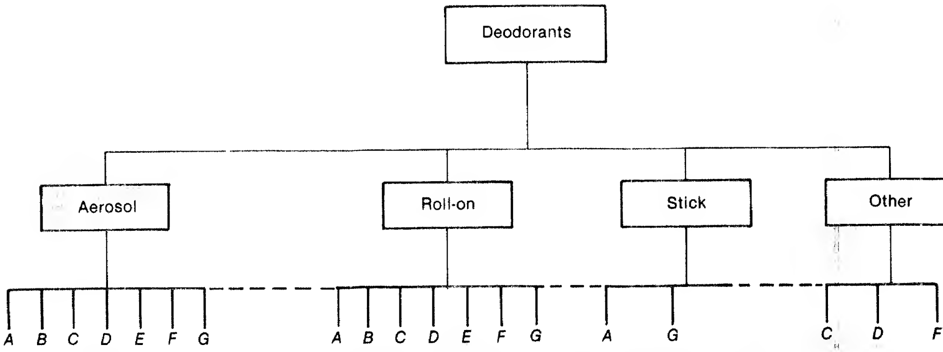


Figure 21–3. Assumed Market Partitioning of the Deodorant Market

An Illustration

Let us now illustrate the par-share analysis and strategic analysis for a deodorant market.[6] We assume that the deodorant market is partitioned first by form and then by brand, as shown in figure 21–3. The problem is to evaluate new-product opportunities under different levels of marketing support. This illustration considers only two marketing decision variables, namely, advertising and price. In all, seven strategy alternatives were tried. These were developed in an ad hoc manner; the intent was to examine the face validity of the PARSHARE model.

The Hendry system forecasts are shown in table 21–1. Given the market partitioning, par-share analysis indicated that a new aerosol brand (alternative 1) would capture a market share of only 2.1 percent, while a new form of deodorant (alternative 2) has the potential to capture a share of 12.6 percent. It should be emphasized that these share forecasts are at equilibrium and assume that the new product's quality and the execution of marketing strategy are at a level comparable to those of competitors. Introducing a new form at a higher price and at optimal advertising (alternative 4) yields a share of 8.5 percent, lower than the expected par of 12.6 percent. However, the return on marketing increases from $11.4 million at par to $15.2 million under alternative 4. Alternative 6 of optimal advertising budget and optimal price with a constraint of reaching at least 10 percent market share results in an advertising budget of $6.5 million and price of $0.96. *Given* a forecast of industry sales and gross margin, this results in a contribution of $14.8 million after accounting for advertising. The effects of various strategies seem to changing in a manner expected a priori.

This analysis can be conducted even before the commencement of new-product development in a corporation. Its main use, however, appears to be as a supplement (or alternative) to pre-test-market estimates.

According to the information of the Hendry Corporation, most new

Table 21-1
Deodorant Market: Strategic Analysis

Alternative	*Advertising ($ Millions)*	*Price (¢/Unit)*	*Market Share (%)*	*Return Marketing*[a] *($, Millions)*
New Aerosol Brand:				
1. Par	1.5	85	2.1	1.9
New Form:				
2. Par	10.7	86	12.6	11.4
3. Par except higher price	10.7	95	11.3	13.0
4. Higher price, optimal[b] advertising	3.2	95	8.5	15.2
5. Optimal[b] advertising and price	3.5	107	7.5	15.8
6. Optimal[b] advertising and price, market share ≥ 10 percent	6.5	96	10.0	14.8
7. Optimal advertising and price, with respect to market share, return ≥ $13 million	8.6	89	11.6	13.0

[a]Return on marketing = gross profits − advertising.
[b]Optimal with respect to return on marketing.

products achieve market shares equal to or close to their par-share estimates. A published Hendry Corporation paper (1976, p. 51) reports that "one manufacturer in making par share determinations on sixteen (16) different products obtained a correlation of .976 between par share predictions made before introduction and actual share of market after introduction."

Discussion

The Hendry system is appropriate as a managerial forecasting technique for frequently purchased consumer products and markets where competitive marketing strategies have a major influence. Some of the disadvantages of the Hendry system are:

1. Since the analyses focus on market share, it is necessary to obtain independent estimates of category sales, perhaps through time-series or regression methods.

2. The models are proprietary, and no comparative studies or extensive testing results have been published. Since the costs are not known, it is necessary for potential users of the system to evaluate the reliability of the estimates and the cost of obtaining them.

The system is parsimonious in its use of input data. This feature is decidedly an advantage, since the availability of data often constrains the choice of methods to be used. The Hendry logic places a greater stress on the model formulation and the assumptions made. Some comparative advantages of the Hendry system are:

1. The system is based on consumer preferences expressed through purchase behavior.
2. The forecasts explicitly account for the structure of the market. This ability strengthens the predictions, since the extent of competition is included in the forecast.
3. The forecasts explicitly account for changes in marketing support and are, therefore, of interest to the manager who would like to know the impact of controllable variables on sales.

Several alternative procedures to forecasting new brand's sales have appeared in the marketing literature. A class of models of interest is primarily concerned with the incidence of purchase events. Three such models are the simple penetration model developed by Fourt and Woodlock (1960), the depth-of-repeat model developed by Eskin (1973), and the STEAM model developed by Massy, Montgomery, and Morrison (1970). The variable of interest in these models may be defined either as the time between purchase events or as the number of purchase events in a given period. Although it is a relatively simple matter to translate one to the other in most models, the latter definition is more relevant for forecasting purposes.

Typically, a purchase-timing model specifies a probability model for the occurrence of purchase events at the level of the individual consumer (or buying unit). The individual purchase frequencies in a given period may be aggregated to give total demand, perhaps by taking into account the heterogeneity of purchase rates among the consumer population. The purchase may refer to a produce class or a specific new product. One of the features of this approach as applied to specific items is that there is no need to define the relevant product class—one merely defines the event of interest to be a purchase of the specific item. This capability is particularly useful when a new product does not fall neatly into traditionally defined product categories, that is, when direct competitors cannot be unambiguously identified.

Note that the Hendry system focuses on brand choice given that a product-class purchase occurs, whereas the penetration models emphasize the

purchase timing or the frequency of purchases in a given period. Traditionally, stochastic models of consumer behavior have followed independent development paths along these dimensions (that is, brand choice and purchase timing). More recently, Zufryden (1977, 1978) has integrated a heterogeneous linear learning model of brand choice with a modified version of the negative binomial distribution for purchase frequency. This composite model is capable of predicting market shares for a given brand over time. (There are other applications of the model that do not concern us here.) The Zufryden model was used to predict a brand's market share over a two-year period beyond an initial period used for parameter estimation. The predictions are quite accurate and compare favorably with alternative penetration models.

No comparisons exist between the predictions made by the Hendry system and other stochastic models appearing in the literature. There exists a clear research need to sort out the distinction between the Hendry approach (for which only some details are known) and other forecasting approaches based on switching behavior. At a more theoretical level, the individual choice behavior models need to be integrated with the aggregate models such as the Hendry system.

Notes

1. In this regard, the ongoing assistance of David Butler of the Hendry Corporation is gratefully acknowledged.

2. Because of the proprietary nature of the Hendry system, technical details available in published literature are necessarily incomeplete. However, some of the conceptual features of the system have made a significant contribution to the formulation of marketing strategy.

3. We must note that market partitioning can be accomplished in several ways. The alternatives depend on the data, models, and techniques used for empirical determination. The Hendry method uses actual behavioral data (that is, switching), an entropy model, and trial and error. In contrast, for example, the method based on consumer perceptions and preferences (for example, Green and Rao 1972) uses similarity/preferential data, models of multidimensional scaling, and several algorithms based on multivariate analytic procedures.

4. This structure is also useful for identifying and evaluating new-product opportunities. In the example of figure 21-1, introducing a new flavor of a given type/brand has less potential than introducing a new brand of a given type.

5. A procedure that deems to subsume several models of market analyses is the minimum-discrimination method (MDI) proposed by

Charnes, Cooper, Learner, and Phillips (1980). The MDI method uses the Kullback-Leibler information theory and derives the models of logit, probit, and Hendry entropy as special cases.

6. The author is indebted to David Butler for demonstrating the analysis in one of his lectures at Cornell University. The basic data are given by Scott Killipps, to whom he is thankful.

References

Bass, Frank M. 1974. "The Theory of Stochastic Preference and Brand Switching." *Journal of Marketing Research* 2 (February):1-20.

Butler, David H. 1975. "Development of Stochastic Marketing Models." Paper presented at the Educational Conference jointly sponsored by TIMS, ORSA, and AMA, Rochester, New York.

Charnes, A.; Cooper, W.W.; Learner, D.B.; and Phillips, F.Y. 1980. "The MDI Method as a Generalization of Logit, Probit and Hendry Analyses in Marketing." Research report CC6368, Center of Cybernetic Studies, University of Texas at Austin.

Day, George S.; Shocker, Allan D.; and Srivatsava, Rajendra. 1979. "Customer-oriented Approaches to Identifying Product-Markets." *Journal of Marketing* 43 (Fall):8-19.

Ehrenberg, A.S.C. 1972. *Repeat-Buying: Theory and Applications.* Amsterdam: North-Holland.

———. 1959. "The Pattern of Consumers' Purchases." *Applied Statistics* 8:26-41.

Ehrenberg, A.S.C., and Goodhardt, G.J. 1968. "The Nature of Brand Switching." *Nature* 220:304.

———. 1969. "A Model of Multi-Brand Buying." *Journal of Marketing Research* 7 (February):77-84.

———. 1979. "The Switching Constant: A Note." *Management Science* 25 (7):703-705.

Eskin, G.J. 1973. "Dynamic Forecasts of New Product Demand Using a Depth of Repeat Model." *Journal of Marketing Research* 10 (May):115-129.

Fourt, L.A., and Woodlock, J.W. 1960. "Early Prediction of Market Success for New Grocery Products." *Journal of Marketing* 25, no. 2 (October):31-38.

Green, Paul E., and Rao, Vithala R. 1972. *Applied Multidimensional Scaling.* Hinsdale, Ill.: Dryden Press.

Hendry Corporation. "Hendro Dynamics: Fundamental Laws of Consumer Dynamics." Chapter 1, 1970; chapter 2, 1971.

———. 1976. "Speaking of Hendry."

Herniter, Jerome D. 1974. "A Comparison of the Entropy Model and the Hendry Model." *Journal of Marketing Research* 11 (February):21–29.

Kalwani, Manohar U. 1979. "The Entropy Concept and the Hendry Partitioning Approach." Working Paper WP1072-79, Sloan School of Management, M.I.T.

Kalwani, Manohar U., and Morrison, Donald G. 1977. "A Parsimonious Description of the Hendry System." *Management Science* 23 (5):467–477.

Massy, W.F.; Montgomery, D.B.; and Morrison, D.G. 1970. *Stochastic Models of Buying Behavior.* Cambridge, Mass.: M.I.T. Press.

Rao, Vithala R., and Sabavala, Darius Jal. 1979. "Methods for Estimation of Market-Structure Parameters." Working Paper, Cornell University.

———. 1978. "Inference of Hierarchical Choice Processes from Analysis of Consumer Panel Data." Working Paper, Cornell University.

Rubinson, Joel R.; Vanhonacker, Wilfried R.; and Bass, Frank M. 1980. "A Note on 'A Parsimonious Description of the Hendry System.' " *Management Science* 26, no. 2 (February):215–226.

Shocker, Alan, and Day, George S. 1976. "Identifying Competitive Product-Market Boundaries: Strategic and Analytical Issues." MSI Working Paper No. 76-112.

Urban, Glen L. 1977. "Hierarchical Definition of Markets." Paper presented at the Joint National ORSA/TIMS Meetings at Atlanta, November 7–9.

Urban, Glen L.; Johnson, Philip L.; and Brudwick, Richard H. 1979. "Market Entry Strategy Formulation: A Hierarchical Modeling and Consumer Measurement Approach." Working Paper No. 1103-80, Sloan School of Management, M.I.T.

Vanhonacker, Wilfried R. 1980. "The Hendry Partitioning Methodology: Theory and Applications." Working Paper No. 288A, Graduate School of Business, Columbia University.

Zufryden, F.S. 1977. "A Composite Heterogeneous Model of Brand Choice and Purchase Timing Behavior." *Management Science* 24 (2):121–136.

———. 1978. "An Empirical Evaluation of a Composite Heterogeneous Model of Brand Choice and Purchase Timing Behavior." *Management Science* 24 (7):761–773.

Appendix 21A: Some Technical Aspects of the Hendry System

In this description, we use the notation of the Hendry Corporation (1970, 1971):

N = number of consumers in product class

N_i = number of consumers buying brand i on any particular occasion

g = number of brands in product class

p_i = market share of brand i

p_{ij} = probability of consumer j choosing brand i

p_{ws} = unconditional probability of switching from brand s

= proportion of consumers switching from brand s on any particular purchase occasion

$p_{(s,s)}$ = unconditional probability of repeat-buying brand s

= proportion of consumer repeat-buying brand s

$p_{(s,h)}$ = unconditional probability of switching from brand s to brand h

$p_{h/s}$ = probability of switching to brand h for a consumer whose previous purchase is brand s and who actually has switched out of brand s

For any individual j,

$$\sum_{i=1}^{i=g} p_{ij} = 1$$

Assume that for a given individual j, the probabilities $\{p_{1j}, p_{2j}, \ldots, p_{gj}\}$ are constant at each purchase occasion. We may term the process *zero-order* (that is, the outcome on a given purchase occasion is not affected by other outcomes) and *stationary* (that is, the probabilities are invariant over time). Different individuals may have probabilities. Then the expected number of individuals choosing brand i is

$$E(N_i) = \sum_{j=1}^{j=N} p_{ij}$$

It is easy to show that the law of detailed balancing is a direct consequence of the zero-order assumption for

$$p_{(s,h)j} = p_{sj}p_{hj} = p_{(h,s)j}$$

Therefore, the expected numbers switching from s to h and from h to s are each equal to

$$\sum_{j=1}^{j=N} p_{sj}\, p_{hj}$$

Now, if equilibrium conditions prevail and if the set of competing items has been correctly identified, then

$$p_{h\,|\,i} = k_i p_h \text{for all } i \neq h$$

From this assertion of proportional switching and by using the law of detailed balancing, it is possible to show that

$$p_{(i,h)} = K_w p_i p_h$$

where K_w is the switching constant and can be expressed as

$$K_w = \frac{\sum_{i=1}^{i=g} p_{wi}}{\sum_{i=1}^{i=g} p_i(1 - p_i)}$$

Note that p_{wi} and p_i refer to the aggregate population and not to a given consumer. And K_w may be interpreted as the ratio of the total proportion switching out of all brands to the total proportion that may be expected to switch under "random" switching behavior. "Random" is defined by the lack of polarization of preferences; that is, all consumers have probability p_i of choosing brand i. However, it is more usually the case that some consumers have developed strong preferences for brand i (with high values of p_{ij}), while others have developed strong preferences for other brands (with low values of p_{ij}). The effect of this heterogeneity is to reduce the amount of switching. Therefore, K_w represents (inversely) the extent to which brand switching is constrained by the polarization of consumer preferences.

Viewed in this sense, K_w may be empirically estimated (with sampling error) from two-period switching data. However, the Hendry system claims

that the level of switching in a partition at equilibrium depends on only the distribution of market shares to the following formula:

$$K_w = \left(\sum_{i=1}^{g} \frac{p_i^2 \ln (1/p_i)}{1 + p_i \ln(1/p_i)} \right) \div \left(\sum_{i=1}^{g} p_i (1 - p_i) \right)$$

The result is shown to be derived from an entropy concept. Different uses of entropy notions by other researchers have generated some controversy. See, for example, Herniter (1974).

Kalwani and Morrison (1977) have shown that if p_i varies over the population according to a beta distribution,

$$f(p_i) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1 - p)^{\beta-1} \qquad 0 < p < 1$$

where $\alpha, \beta > 0$, then

$$K_w = \frac{\alpha + \beta}{1 + \alpha + \beta}$$

It should be noted that K_w when computed in this way requires data on at least two choice occasions in order for α and β to be estimated. A point of view that raises a distinction between theoretical and empirical switching is presented by Ehrenberg and Goodhardt (1979).

The hierarchical market structure or partitioning scheme of the Hendry system assumes that consumer preferences divide the total set of potentially competing items into two partitions. For example, in figure 21-1, partitions at the highest level of switching are defined by type and then brand. Within each type/brand partition, the laws of switching constants K_{fa}, K_{fb}, . . . , K_{fm}, K_{fn} may characterize the switching for *flavor* alternatives under brands *A*, *B*, . . . , *M*, *N*. In addition, there is some switching between these partitions. The Hendry system posits that at each level of the hierarchy the switching laws will apply, but with different values of K_w. In our example, the switching between the brands for a given type may be characterized by K_{b1}, K_{b2}, K_{b3}, . . . for brands under types 1, 2, 3, ... Finally, the switching between types may be characterized by K_t. If a partition has been correctly identified, then the observed switching should correspond to the theoretical switching computed using the estimated K_w's.

We believe that the Hendry methodology identifies a partition in a trial-and-error fashion. That is, a hypothesized partitioning structure is examined to see if the K_w estimates reproduce the observed switching matrix. One

possible criterion for selecting a partition is the closeness of fit of the expected interitem switching patterns to actual switching data. The expected interitem switching may be computed as follows. Let us consider a simple version of the structure in figure 21-1. Let

P_{tbf} = market shares for flavor f within given type t and brand b

P_{tb} = market shares for brands b within a given type

P_t = market shares of types t

K_{tb} = switching constant across flavors for given type/brand

K_t = switching constant across brands for given type

K = switching constant across types

N = total market of consumers

The shares are normalized at each level, so that

$$\Sigma_f P_{tbf} = 1$$
$$\Sigma_b P_{tb} = 1$$
$$\Sigma_t P_t = 1$$

First, consider switching between flavors 1 and 2 of brand A, type 1. The proportional switching is given by

$$P_{(1A1,\,1A2)} = K_{1A} P_{1A1} P_{1A2}$$

The expected number of consumers switching is

$$N_{(1A1,\,1A2)} = (NP_1 P_A)(K_{1A} P_{1A1} P_{1A2})$$

Similarly,

$$N_{(2D1,\,2D3)} = (NP_2 P_{2D})(K_{2D} P_{2D1} P_{2D3})$$

Next, the switching between two brands A and B is

$$P_{(1A,\,1B)} = K_1 P_{1A} P_{1B}$$
$$N_{(1A,\,1B)} = (NP_1)(K_1 P_{1A} P_{1B})$$

Because of proportional switching, the switching between flavor 1, brand A, and flavor 1, brand B, is

$$N_{(1A1,\,1B1)} = P_{1A1}P_{1B1}(NP_1)(K_1P_{1A}P_{1B})$$

Finally, consider switching between types 1 and 2 as given by

$$P_{(1,2)} = KP_1P_2$$

$$N_{(1,2)} = (N)(KP_1,P_2)$$

The switching between brand A, type 1, and brand D, type 2, is

$$N_{(1A,\,2D)} = P_{1A}P_{2D}(N)(KP_1P_2)$$

The switching between flavor 1, brand A, and flavor 3, brand D, is

$$N_{(1A1,\,2D3)} = P_{1A1}\,P_{2D3}(P_{1A}P_{2D})(N)(KP_1P_2)$$

The switching and partitioning for a product category remains stable when the market is in equilibrium. When a new product enters the market, a new equilibrium is achieved after some transition period. As new brands enter a market, the level of switching tends to decrease as, presumably, consumer needs are more precisely met and the polarization of preferences increases. Technical details of how the new product develops customers are not known. The following two concepts are implicitly used in the Hendry analysis, however:

1. The consumers most likely to try and to accept the new product are those who do not have strong preferences for existing brands.
2. The potential number of customers who will develop strong preferences for the new brand is directly related to the level of switching. A higher level of switching indicates greater vulnerability.

The Hendry system is able to estimate the equilibrium market share for a new product given the market structure and switching parameters and by assuming the new product has "par" quality and "par" marketing support. ("Par" is defined by the existing products in the category.) This output is called the par share. This forecast of a potential new product's sales is made solely on the basis of the existing products.

Response functions to advertising, promotion, price, and distribution are incorporated in the system. The effects of some strategies are completely "reversible." That is, market-share gains are not permanent. The effects of some other strategies are used to estimate market share for existing or new products under alternative marketing strategies. In addition, it is possible to find strategies (if feasible) to meet certain requirements of market share and financial return.

References

Ehrenberg, A.S.C., and Goodhardt, G.J. 1979. "The Switching Constant: A Note." *Management Science* 25 (7):703–705.

Hendry Corporation. "Hendro Dynamics: Fundamental Laws of Consumer Dynamics." Chapter 1, 1970; chapter 2, 1971.

Herniter, Jerome D. 1974. "A Comparison of the Entropy Model and the Hendry Model." *Journal of Marketing Research* 11 (February);21–29.

Kalwani, Manohar U., and Morrison, Donald G. 1977. "A Parsimonious Description of the Hendry System." *Management Science* 23 (5):467–477.

Appendix 21B: Hendrodynamics Strategic Analysis of a New-Brand Introduction

David Butler

SECTION 1: Inputting Data
Calculation of PAR Profile
Calculaton of Critical Budget
Determining Share & Profit Response to Advertising and Distribution Changes

SECTION 2: Determining Optimal Advertising and Price Combinations for the New Brand

Section 1.

1 *Data Input:*

LEVEL: Level in the Category Partitioning Structure—This is a simple illustration—All brands are at level 1, i.e. they all compete directly with one another

VMS: Category Volume × brand unit margin × brand share of market

M: brand unit margin equals unit price to trade after promotional allowances (PT) minus unit variable cost (manufacturing, selling, and distribution)

E: brand advertising expenditure

$: brand price promotion allowance

S: brand share of market

P: brand price to consumer

PAVG: category average price to consumer

COST: unit variable cost

PT: unit price to trade after promotional allowances

2 *Data Input:*

D: brand distribution levl; %ACV minus out of stock

DSUM: sum of brand distribution levels

3 *Starting Conditions:*

Return on Marketing = VMS − E

4 Program is told to compute Par Profile for the next brand into the market. Analysis assumes the new brand is neither superior to or inferior to competition.

5 Par Share potential for a new brand entry at par advertising, promotion, distribution, and price to consumer

In this example, there is good share and profit (return on marketing) potential for a new brand entry. The potential new brand is alternative 5.

6 The critical (minimum) advertising budget and its allocation by introductory quarter. Budget less than the critical result in long term share of market damage.

7 If the new brand spends the critical budget and is priced at Par (100 price to consumer) and has Par Distribution (89%) then the expected introductory year ending share of market is 15.05%.

8 New brand expected share of market and return on marketing response to advertising and distribution.

```
*RUN "HD

                  HENDRODYNAMICS - STRATEGIC ANALYSIS
DESCRIPTION  NEW BRAND INTRODUCTION

PROGRAM TO BE SAVED AUTOMATICALLY AFTER EACH CHARGE UNDER ? WIND
ALT(   1) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"
 ? "1,39680,8500,9920,62,100,100,40,80"
ALT(   2) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"
 ? "1,13440,3350,3360,21,100,100,40,80"
ALT(   3) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"          (1)
 ? "1,5760,1920,1440,9,100,100,40,80"
ALT(   4) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"
 ? "1,5120,1710,1280,8,100,100,40,80"
ALT(   5) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"
 ? EDIT
(1=LIST,2=INPUT REMAINING,3=CORRECT,4=DELETE,5=ADD) ALTERNATIVE(S) ? 1
ALT  LEVEL     VMS      E       $ SHARE  PRICE   PAVG   COST    PT  (M/MS)
 1)    1.0   39680    8500   9920 62.00 100.00 100.00  40.00  80.00  1.00
 2)    1.0   13440    3350   3360 21.00 100.00 100.00  40.00  80.00  1.00
 3)    1.0    5760    1920   1440  9.00 100.00 100.00  40.00  80.00  1.00
 4)    1.0    5120    1710   1280  8.00 100.00 100.00  40.00  80.00  1.00
(1=LIST,2=INPUT REMAINING,3=CORRECT,4=DELETE,5=ADD) ALTERNATIVE(S) ? 2
ALT(   5) INPUT:"LEVEL,VMS,E,$,S,P,PAVG,C,PT,(M/MS)"
 ? NO
"DISPLAY ALTERNATIVES" ? "1,2,3,4
CORE IMAGE SAVED UNDER WIND
MOVE( 1 ) INPUT ALT,E,P ? 92,0,0
INPUT: D, DSUM, %D+, CD
ALT(  1):  ? 99,308,0,0
ALT(  2):  ? 80,308,0,0                                          (2)
ALT(  3):  ? 66,308,0,0
ALT(  4):  ? 63,308,0,0
CORE IMAGE SAVED UNDER WIND
MOVE( 1 ) INPUT ALT,E,P,D ? 0,0,0,0
```

```
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500        9920        0  99.0    100.00     62.00      31180

  2)    3350        3360        0  80.0    100.00     21.00      10090

  3)    1920        1440        0  66.0    100.00      9.00       3840

  4)    1710        1280        0  63.0    100.00      8.00       3410
MOVE( 1 ) INPUT ALT,E,P,D ? 100,1,0,0
CORE IMAGE SAVED UNDER WIND
INPUT ALT,LEV DESIGNATOR,MX DESIGNATOR 5,1,0
ALTERNATIVES HAVE BEEN RENUMBERED
PAR PROFILE
LEVEL      VMS       E        $       S        P   PAVG       C      PT
1          8039     2593     2010 12.56 100.00 100.00   50.00 100.00
D 89.16  DSUM 308        %D+ 0          CD 0
"DISPLAY ALTERNATIVES" ? "1,2,3,4,5"
CORE IMAGE SAVED UNDER WIND
MOVE( 1 ) INPUT ALT,E,P,D ? 0,0,0,0
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500        8674        0  99.0    100.00     54.21      26196

  2)    3350        2938        0  80.0    100.00     18.36       8402

  3)    1920        1259        0  66.0    100.00      7.87       3116

  4)    1710        1119        0  63.0    100.00      7.00       2767

  5)    2593        2010        0  89.2    100.00     12.56       5446
MOVE( 1 ) INPUT ALT,E,P,D ?
```

(3)

(4)

(5)

```
*LOAD "WIND
*RUN 700


MOVE( 1 ) INPUT ALT,E,P,D ? 101,0,0,0
ALT, % OF ALT TO POLARIZE, PURCHASE CYCLE IN MONTHS 5,100,3
CORE IMAGE SAVED UNDER WIND
CRITICAL BUDGET PERIOD IN MONTHS (3=QUARTER) ? 3
INPUT E,P,D ? 0,100,89
E 0 P 100 D 89 S 8.57734 R 5489.49

CRITICAL BUDGET
PERIOD 1   P ? 100 D ? 89    ECRIT   2566 SHARE 16.72 FR COMP 0.000
PERIOD 2   P ? 100 D ? 89    ECRIT   1889 SHARE 15.78 FR COMP 0.264
PERIOD 3   P ? 100 D ? 89    ECRIT   1044 SHARE 13.91 FR COMP 0.593
PERIOD 4   P ? 100 D ? 89    ECRIT    513 SHARE 11.94 FR COMP 0.800
FOUR PERIOD TOTALS   E 6012.2 S 9.96222 R 363.613

INPUT BUDGET FOR FIRST 8 PERIODS
 ? 0,0,0,0,0,0,0,0
MOVE( 1 ) INPUT ALT,E,P,D ? 5,6012,100,89
MOVE( 2 ) INPUT ALT,E,P,D ? 0,0,0,0
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500        8427        0  99.0    100.00     52.67      25207
```

(6)

```
  2)    3350        2854        0  80.0     100.00      17.84       8067
  3)    1920        1223        0  66.0     100.00       7.65       2973
  4)    1710        1087        0  63.0     100.00       6.80       2639
  5)    6012        2408        0  89.0     100.00      15.05       3621   (7)
MOVE( 1 ) INPUT ALT,E,P,D ? 90,1,1,0
"DISPLAY ALTERNATIVES" ? "5"

MOVE( 1 ) INPUT ALT,E,P,D ? 95,1,0,0

INPUT NUMBER OF DRIVERS ? 1

DRIVER ( 1 ):
INPUT ALT.#  ? 5
INPUT START VALUE,END VALUE,INCREMENT OF E
 ? 0,6000,1000
INPUT START VALUE,END VALUE,INCREMENT OF P
 ? 100,100,1
INPUT START VALUE,END VALUE,INCREMENT OF D
 ? 60,90,10

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON   (8)
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)      0         900      0  60.0    100.00       5.62       3600

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)      0        1064      0  70.0    100.00       6.65       4255

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)      0        1227      0  80.0    100.00       7.67       4907

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)      0        1388      0  90.0    100.00       8.68       5554

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)   1000        1109      0  60.0    100.00       6.93       3438

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)   1000        1308      0  70.0    100.00       8.18       4233

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)   1000        1505      0  80.0    100.00       9.41       5019

 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  5)   1000        1699      0  90.0    100.00      10.62       5796
```

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 2000	1262	0 60.0	100.00	7.89	3048

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 2000	1485	0 70.0	100.00	9.28	3942

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 2000	1706	0 80.0	100.00	10.66	4823

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 2000	1922	0 90.0	100.00	12.01	5689

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 3000	1380	0 60.0	100.00	8.63	2520

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 3000	1622	0 70.0	100.00	10.14	3488

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 3000	1860	0 80.0	100.00	11.62	4440

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 3000	2093	0 90.0	100.00	13.08	5373

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 4000	1475	0 60.0	100.00	9.22	1899

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 4000	1731	0 70.0	100.00	10.82	2925

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 4000	1983	0 80.0	100.00	12.39	3932

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 4000	2229	0 90.0	100.00	13.93	4917

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 5000	1553	0 60.0	100.00	9.70	1211

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 5000	1821	0 70.0	100.00	11.38	2284

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 5000	2084	0 80.0	100.00	13.02	3336

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 5000	2341	0 90.0	100.00	14.63	4362

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 6000	1618	0 60.0	100.00	10.11	472

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 6000	1896	0 70.0	100.00	11.85	1585

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 6000	2168	0 80.0	100.00	13.55	2673

ADVERTISING EXPENDITURE	PROMOTION ALLOWANCE	DISTRIBUTION IC COST DIST	PRICE TO CONSUMER	SHARE OF MARKET	RETURN ON MARKETING
5) 6000	2434	0 90.0	100.00	15.21	3734

Section 2

1. Alternative 5 is the expected or Par Profile for the next new brand in the market.
2. Allocation of advertising and promotion which optimizes profit for alternative 5 while maintaining share at 12.5%. There is little profit opportunity going from 1 to 2 i.e., the Par profile is close to optimum profit for a 12.5 share.
3. Optimum profit advertising and price promotion allocation for alternative 5 while maintaining a 14% share of market.
4. Same as 3 except it's for a 15% share.
5. Same as 3 except it's for a 16% share.
6. Maximum profit advertising budget—price to consumer for alt 5 held constant.
7. Maximum profit advertising and price combination.
8. Combinations of advertising and price promotions which have the same marginal returns.

A marginal return (MR) of 1.00 is profit maximization. The last dollar of advertising or promotion generating $1.00 of contribution to profit.

MR)p = Marginal Return of Advertising, brand price held constant.
MR)$ = Marginal Return of Advertising, brand price promotion budget held constant.
MR)E = Marginal Return of price promotion, brand advertising budget held constant.

```
*LOAD "WIND
*RUN 700
MOVE( 1 ) INPUT ALT,E,P,D ? 0,0,0,0
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500       8674      0  99.0     100.00    54.21     26196

  2)    3350       2938      0  80.0     100.00    18.36      8402

  3)    1920       1259      0  66.0     100.00     7.87      3116

  4)    1710       1119      0  63.0     100.00     7.00      2767

  5)    2593       2010      0  89.2     100.00    12.56      5446   (1)
MOVE( 1 ) INPUT ALT,E,P,D ? 93,0,0,0
HD ALLOCATION PACKAGE
NO. OF ALTS CONSTRAINED OR MOVED IN OPTION ? 1
CRITERION (1=MAX R,2=MAX S,3=MIN R,4=MIN S) 1
ENTER ALT, NO. OF CONSTRAINTS & THE CONSTRAINTS FOR EACH ALT
CONSTRAINTS MAY BE PLACED ON S,R,$,E OR B.
END WITH ALTERNATIVE TO BE OPTIMIZED.
ALT, NO. OF CONSTRAINTS ? 5,1
CONSTRAINT 1 ,MIN VALUE ,MAX VALUE ? S,12.5,100
INPUT STARTING POINT
ALT( 5) INPUT E,P,D  ? 2600,100,89.2
INPUT SECONDARY POINTS
ALT( 5) INPUT E,P  ? 4000,95
INPUT DATA ACCEPTED
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500       8680      0  99.0     100.00    54.25     26219

  2)    3350       2940      0  80.0     100.00    18.37      8410

  3)    1920       1260      0  66.0     100.00     7.87      3120

  4)    1710       1120      0  63.0     100.00     7.00      2770

  5)    3317       1096      0  89.2     105.65    12.50      5589   (2)

CORE IMAGE SAVED UNDER WIND
MOVE( 1 ) INPUT ALT,E,P,D ? 93,0,0,0
HD ALLOCATION PACKAGE
NO. OF ALTS CONSTRAINED OR MOVED IN OPTION ? 1
CRITERION (1=MAX R,2=MAX S,3=MIN R,4=MIN S) 1
ENTER ALT, NO. OF CONSTRAINTS & THE CONSTRAINTS FOR EACH ALT.
CONSTRAINTS MAY BE PLACED ON S,R,$,E OR B.
END WITH ALTERNATIVE TO BE OPTIMIZED.
ALT, NO. OF CONSTRAINTS ? 5,1
CONSTRAINT 1 ,MIN VALUE ,MAX VALUE ? S,14,100
```

```
INPUT STARTING POINT
ALT( 5) INPUT E,P,D  ? 4000,95,89
INPUT SECONDARY POINTS
ALT( 5) INPUT:E,P  ? 5000,90
INPUT DATA ACCEPTED
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500       8531        0  99.0    100.00     53.32     25625

  2)    3350       2890        0  80.0    100.00     18.06      8208

  3)    1920       1238        0  66.0    100.00      7.74      3034

  4)    1710       1101        0  63.0    100.00      6.88      2693

  5)    3878       2642        0  89.0     97.76     14.00      4680    3
MOVE( 1 ) INPUT ALT,E,P,D ? 93,0,0,0
HD ALLOCATION PACKAGE
NO. OF ALTS CONSTRAINED OR MOVED IN OPTION ? 1
CRITERION (1=MAX R,2=MAX S,3=MIN R,4=MIN S) 1
ENTER ALT, NO. OF CONSTRAINTS & THE CONSTRAINTS FOR EACH ALT
CONSTRAINTS MAY BE PLACED ON S,R,$,E OR B.
END WITH ALTERNATIVE TO BE OPTIMIZED.
ALT, NO. OF CONSTRAINTS ? 5,1
CONSTRAINT 1 ,MIN VALUE ,MAX VALUE ? S,15,100
INPUT STARTING POINT
ALT( 5) INPUT E,P,D  ? 4000,90,89
INPUT SECONDARY POINTS
ALT( 5) INPUT:E,P  ? 5000,85

INPUT DATA ACCEPTED
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500       8432        0  99.0    100.00     52.70     25228

  2)    3350       2856        0  80.0    100.00     17.85      8074

  3)    1920       1224        0  66.0    100.00      7.65      2976

  4)    1710       1088        0  63.0    100.00      6.80      2642

  5)    3934       4041        0  89.0     91.45     15.00      4025    4
MOVE( 1 ) INPUT ALT,E,P,D ? 93,0,0,0
HD ALLOCATION PACKAGE
NO. OF ALTS CONSTRAINED OR MOVED IN OPTION ? 1
CRITERION (1=MAX R,2=MAX S,3=MIN R,4=MIN S) 1
ENTER ALT, NO. OF CONSTRAINTS & THE CONSTRAINTS FOR EACH ALT.
CONSTRAINTS MAY BE PLACED ON S,R,$,E OR B.
END WITH ALTERNATIVE TO BE OPTIMIZED.
ALT, NO. OF CONSTRAINTS ? 5,1
CONSTRAINT 1 ,MIN VALUE ,MAX VALUE ? S,16,100
INPUT STARTING POINT
ALT( 5) INPUT E,P,D  ? 5000,80,89
INPUT SECONDARY POINTS
ALT( 5) INPUT E,P  ? 6000,75
INPUT DATA ACCEPTED
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500       8333        0  99.0    100.00     52.08     24831

  2)    3350       2822        0  80.0    100.00     17.64      7939

  3)    1920       1210        0  66.0    100.00      7.56      2918
```

```
4)     1710        1075       0  63.0    100.00      6.72      2591

5)     3988        5466       0  89.0     85.81     16.00      3348    (5)

MOVE( 1 ) INPUT ALT,E,P,D ? 0,0,0,0
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)     8500        8674       0  99.0    100.00     54.21     26196

  2)     3350        2938       0  80.0    100.00     18.36      8402

  3)     1920        1259       0  66.0    100.00      7.87      3116

  4)     1710        1119       0  63.0    100.00      7.00      2767

  5)     2593        2010       0  89.2    100.00     12.56      5446
MOVE( 1 ) INPUT ALT,E,P,D ?
MOVE( 1 ) INPUT ALT,E,P,D ? 98,0,0,0
INPUT ALT,CRITERION ? 5,P
DO YOU WISH TO GENERATE A PROFILE ? NO
INPUT MR)P ? 1.00

MR)P=  1.0000      MR)$=  1.1073      MR)E=   .5708
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     1080        1703 100.00  10.64      5732                (6)
MOVE( 1 ) INPUT ALT,E,P,D ? 98,0,0,0
INPUT ALT,CRITERION ? 5,A
DO YOU WISH TO GENERATE A PROFILE ? NO
INPUT MR)A ? 1.00
E=     891 P=124.95 MR)$= 1.649 MR)E= 0.577
E=    1414 P=126.49 MR)$= 1.204 MR)E= 0.969
E=    1529 P=126.24 MR)$= 1.034 MR)E= 1.005
E=    1526 P=126.34 MR)$= 0.999 MR)E= 0.998

MR)P=  1.0004      MR)$=  1.0005      MR)E=  1.0005
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     1526       -1472 126.34   8.31      6596                (7)
MOVE( 1 ) INPUT ALT,E,P,D ? 98,0,0,0
INPUT ALT,CRITERION ? 5,A
DO YOU WISH TO GENERATE A PROFILE ? YES
INPUT STARTING VALUE, END VALUE, INCREMENT OF E ? 2000,4000,500

MR)P=   .8727      MR)$=  0.8589      MR)E=   .8590
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     2000        -950 120.35   9.46      6520

MR)P=   .7546      MR)$=  0.7489      MR)E=   .7487
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     2500        -264 114.44  10.65      6283

MR)P=   .6506      MR)$=  0.6640      MR)E=   .6639
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN                 (8)
  5)     3000         545 108.89  11.80      5896

MR)P=   .5592      MR)$=  0.5964      MR)E=   .5963
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     3500        1462 103.66  12.91      5369

MR)P=   .3747      MR)$=  0.5313      MR)E=   .5319
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)     4000        6246  83.02  16.55      2996
```

```
MOVE( 1 ) INPUT ALT,E,P,D ? 98,0,0,0
INPUT ALT,CRITERION ? 5,A
DO YOU WISH TO GENERATE A PROFILE ? YES
INPUT STARTING VALUE, END VALUE, INCREMENT OF E ? 3500,4000,100

MR)P=    .5592      MR)$=  0.5964      MR)E=    .5963
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    3500        1462 103.66  12.91       5369

MR)P=    .5423      MR)$=  0.5845      MR)E=    .5845
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    3600        1656 102.65  13.13       5249

MR)P=    .5257      MR)$=  0.5729      MR)E=    .5729
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    3700        1855 101.64  13.35       5123

MR)P=    .5078      MR)$=  0.5618      MR)E=    .5608
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    3800        2122 100.32  13.61       4964

MR)P=    .4504      MR)$=  0.5466      MR)E=   ..5475
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    3900        3806  92.46  14.84       4164

MR)P=    .3747      MR)$=  0.5313      MR)E=    .5319
 ADVERTISING  PROMOTION  PRICE  SHARE     RETURN
  5)    4000        6246  83.02  16.55       2996
MOVE( 1 ) INPUT ALT,E,P,D ? 0,0,0,0
 ADVERTISING  PROMOTION  DISTRIBUTION  PRICE TO  SHARE OF  RETURN ON
 EXPENDITURE  ALLOWANCE  IC COST DIST  CONSUMER  MARKET    MARKETING

  1)    8500        8674        0  99.0    100.00     54.21     26196

  2)    3350        2938        0  80.0    100.00     18.36      8402

  3)    1920        1259        0  66.0    100.00      7.87      3116

  4)    1710        1119        0  63.0    100.00      7.00      2767

  5)    2593        2010        0  89.2    100.00     12.56      5446
MOVE( 1 ) INPUT ALT,E,P,D ?
```

22 Applications of Diffusion Models: Some Empirical Results

Kenneth D. Lawrence and
William H. Lawton

The product life cycle is an important concept in product planning since it shows a likely future performance of the product in terms of both sales and profit. Moreover, it also provides a marketing-oriented basis for product taxonomy. In such a product taxonomy, the products are classified into homogeneous groups which react in a similar fashion to marketing stimuli. The value of the product life cycle lies in its ability to be a forecasting tool which exhibits a sales-and-profit profile of the product.

If the product life-cycle curve were known prior to marketing a product and the market was stable and unchanging, the manufacturing and marketing decisions concerning the product would be greatly simplified. Unfortunately, the product life-cycle curve is known only after the death of the product, too late to be of use in early manufacturing and marketing decisions. Also, the marketplace is subject to considerable change introduced by the firm's introducing the product at hand, as well as the competitors of that firm. This situation has given rise to a great deal of effort aimed at predicting the critical product life cycle prior to product introduction (in fact, prior to manufacturing commitment), using methods which allow one to manipulate assumptions about the ever-changing marketplace into which the product is to be thrust (Bass 1969; Davies 1977; Midgley 1977; Mahajan and Muller 1979; Geurts and Reinmuth 1980).

In applying the product life-cycle model to the area of new-product planning and forecasting, the vast majority of the applications have been to consumer products. However, the focus of this chapter is directed toward the development of product life-cycle curves, before there is actual sales-data history, for both the industrial and the consumer products.

The Model

The problem of predicting life-cycle curves prior to the existence of any sales data has two distinct components. The first involves the construction of a simple, analytic model which will allow one to generate an entire product life cycle from very few interpretable parameters such as market size. With such a model, the problem of predicting the entire product life cycle is

reduced to that of predicting two or three meaningful parameters. The model suggested in this chapter arises in epidemiology and meets the above-mentioned criterion.

The second component deals with methods of estimating these key parameters and their changes with time prior to the existence of any sales data. One must, of course, always remember that all mathematical models are suspect. No matter how simple or appealing a model may be, the model and its parameter estimates cannot be trusted until it can be shown that the model is consistent with, and provides an explanation for, a substantial history of earlier new-product life cycles. This model has been applied to several product classes of industrial products, and it has consistently provided an excellent representation of first-purchase data (Lawton and Lawton 1979). The model formulation is

$$\frac{dS(t)}{dt} = p^* \frac{N_0 + S(t)}{N + N_0} [N - S(t)] \tag{22.1}$$

or

$$S(t) = \frac{N + N_0}{(1 + (N/N_0)e^{-p^*t})} - N_0 \tag{22.2}$$

where $S(t)$ = cumulative unit sales of product up to time t

N_0 = "effective" numbr of prior product users

N = market size

p^* = rate parameter

It should be noted that if we set

$$A = p^* \frac{N_0}{N + N_0}$$

and

$$B = \frac{p^*}{N + N_0}$$

equation 22.1 reduces to

$$\frac{dS(t)}{dt} = A[(N - S(t)] + BS(t)[N - S(t)] \tag{22.3}$$

Equation 22.3 is essentially the formulation for the Bass-type diffusion models, where A is the coefficient of innovation and B is the coefficient of imitation (Bass 1969). The suggested formulation, equation 22.1, assumes that $A/B = N_0$.

The utilization of equation 22.1, or equation 22.2, involves the estimation of N, p^*, and N_0. If the time t is in years, the sales in year t are simply given by $S(t) - S(t - 1)$. Given annual sales data, one can seek those parameters N, p^*, and N_0 which give the best agreement between the observed sales in year t and the predicted sales $S(t) - S(t - 1)$ in year t. Iterative methods for finding these "best" parameters are readily available and involve the use of nonlinear least-squares techniques such as those described in Draper and Smith (1967). Other authors have recommended using linear least-squares techniques to fit the difference equation corresponding to equation 22.1. Such a linearized approach violates the assumptions of the error structure in linear regression and will lead to poor fits and highly biased estimates. The model, represented by equation 22.2, has been fitted to more than thirty sets of historical data on new-product life cycles, and the agreement has been excellent.

If we are now faced with a new-product introduction and wish to use the model as a means of forecasting the timing of sales, then we must estimate the three key parameters, N, p^*, and N_0. Once we have estimated these three values, we can plot product sales as a function of time for the entire cycle. It is, however, clear that the overall market size N is dependent on pricing, advertising, and distribution as well as on inherent product features. For a given market strategy and a static marketplace without changes in competition, N has some "true' but unknown value. Direct market survey or other market-research techniques will allow us, at least conceptually, to estimate N prior to product introduction. But what about the estimation of the remaining two parameters?

The "effective" number of prior users at time $t = 0$, N_0, created through advertising and distribution of free samples, is more difficult to interpret and hence to estimate a priori. We can, however, introduce a more easily interpreted substitute parameter by means of a mathematical trick. Given N and p^*, the value of N_0 is completely determined by the number of sales in the first year. Using equation 22.2 with $t = 1$ and solving for N_0, one has

$$N_0 = \frac{NS_1e^{-p^*}}{N(1 - e^{-p^*}) - S_1} \tag{22.4}$$

where $S_1 = S(1)$, the sales in year 1. We know how we can go about producing an estimate of N, and the same sort of direct market-research techniques will allow us to develop an estimate of S_1. Hence, if we can now find a

method for estimating p^*, the parameter N_0 can be obtained from 22.4. Our basic model can then be viewed either as a function of N, p^*, N_0 or as a function of N, p^*, S_1.

Empirical Findings

In our examination of the thirty-odd historical new-product life cycles, we obtained estimates of the rate parameter p^*. Figure 22–1 shows a plot of the various values of p^* for twelve of these products. An interesting structure emerges. For the nine consumer products, all but one of the values of p^* fall between 0.4 and 0.55 (the exception being color TV). The average value for this grouping is about $p^* = 0.5$. Similarly, the values of p^* for the three commercial products tend to be grouped together. The commercial products have p^* between 0.65 and 0.68 with an average value near $p^* = 0.66$.

These historic data suggest an obvious approach to the estimation of the remaining parameter p^*. For commercial products, simply take an a priori estimate of $p^* = 0.66$; and for consumer products, take $p^* = 0.5$. One would, of course, like to verify that restricting p^* by such a rule does not seriously degrade the model's ability to represent the new-product sales data. As a rough cut at this verification, we refitted the original sales data, using the model with p^* restricted to 0.5 or 0.66 depending on product type. Figure 22–2 shows a plot of the predicted peak-sales level before and after constraints on p^*. Clearly there is no significant degradation. Figure 22–3 shows the same type of plot for time to peak sales. While figure 22–3 shows some loss of precision, it is within acceptable limits.

We then developed an approach to estimating the life cycle for a new product in a stable marketplace. Given the market plan including features, price, advertising, distribution, and so on, take p^* as 0.5 for consumer products and 0.66 for commercial products, estimate N and S_1 by traditional market-research methods, obtain N_0 by 22.4, and generate the life cycle by 22.2. Clearly, if one is dealing with a narrow product class and has prior sales histories of similar products, then one may be able to develop and use a better a priori estimate of p^* than that given by our 0.5-or-0.66 rule.

The assumption that p^* is essentially constant over time is fairly reasonable. However, the assumption that the market size N will remain constant during the life of the product seems quite unreasonable. At various times in the life of the product, changing prices and introduction of competitive products would have significant impact on the market size remaining for the product. One of the unique features of this epidemiologic model is its ability to handle such dynamic changes in market size due to external variables.

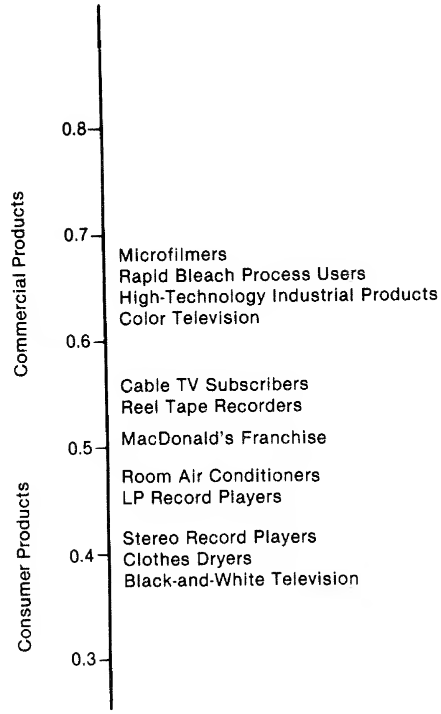


Figure 22–1. Comparison of Rate Constants for Consumer and Commercial Products

The epidemiology-based model has what we might call a conditional nature. That is, at any point in time we can stop the model and then restart it, based on current conditions. To see this capability, let us see how one would use the model to develop a market analysis for a new product.

Assume that through extensive market research we have estimated that our new product will have a market size of N = 100,000 units if introduced into a static current marketplace under our marketing base plan A. We are planning to introduce the product in 1980 and have estimated, again through market research, the first year's sales $S(1)$ to be 5,000 units. Since we have a commercial product, we take $p^* = 0.6$. The "effective" number of prior users in 1980 is, by equation 22.4, just N_0 = 5,967. The resulting life cycle is shown in figure 22–4, and this would be the sales forecast under plan A in a static, unchanging marketplace.

Let us further assume that by 1986 we expect to be able to offer an

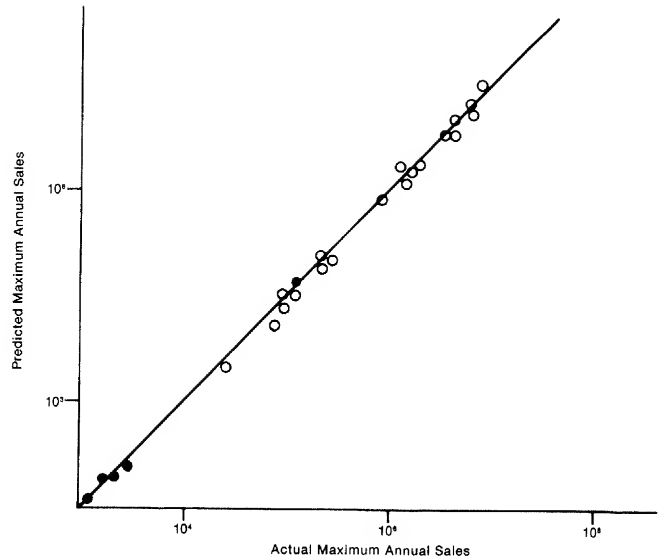


Figure 22-2. Plot of Predicted Peak-Sales Level before and after Constraint on p^*

improved product with a new feature. Now what is our sales forecast? The sales between 1980 and 1986 are, of course, the same as in figure 22-4. By 1986, the model tells us that we would have sold 47,000 units, leaving a remaining market size of 53,000 units for potential sales. Our first question in developing the sales forecast for 1986 and beyond deals with how the new-product feature affects the remaining post-1986 market size.

If the added feature is not perceived new by the marketplace, then the post-1986 market size is just the 53,000 units remaining from the original market. If we now restart the model in 1986 using the effective-number-of-prior users form of the model with the now-current 1986 conditions, we would have N = 53,000 and p^* = 0.66. The effective number of users N_0 would, in 1986, be the 5,967 effective number or prior users in 1980, determined from equation 22.4, plus the 47,000 users who have bought between 1980 and 1986. Thus, N_0 = 5,967 + 47,000 = 52,967. Using these values for the model parameters produces a forecast for 1986, which exactly coincides with the original forecast in figure 22-4. Thus, even though the forecast has been done in two pieces, we have results identical to the one-piece forecast of figure 22-4.

On the other hand, let us assume that through market research we have estimated that the new feature in 1986 will increase the market for the prod-

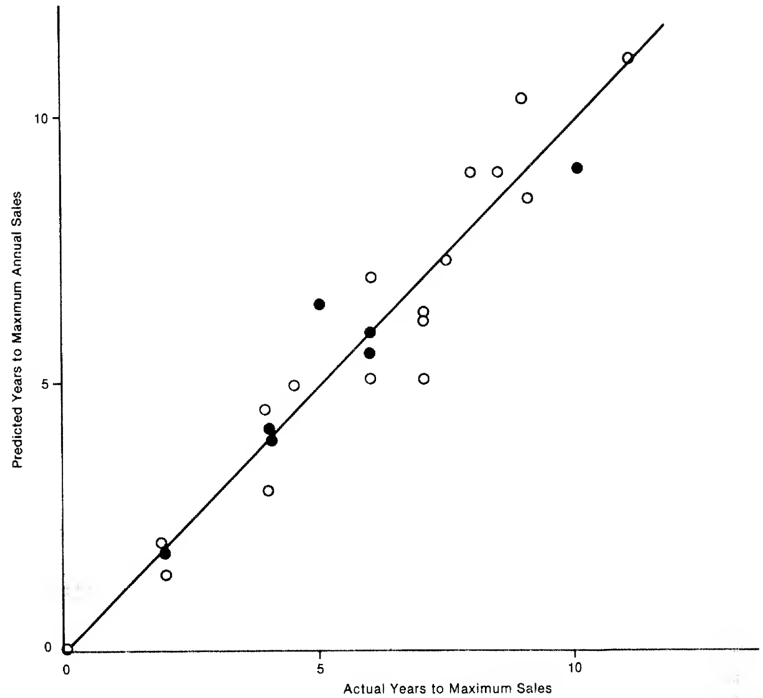


Figure 22–3. Plot of Predicted Time to Peak Sales before and after Constraint on p^*

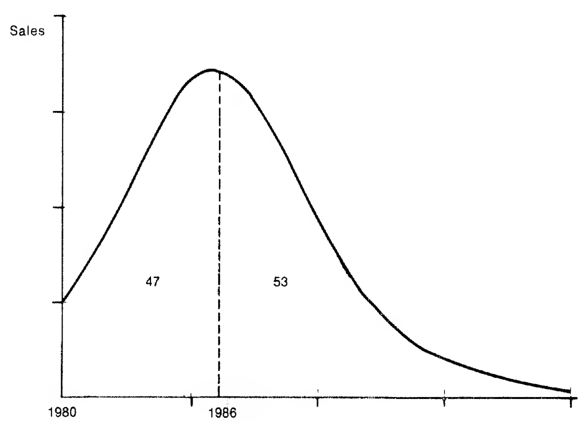


Figure 22–4. Sales Forecast, 1980–1986

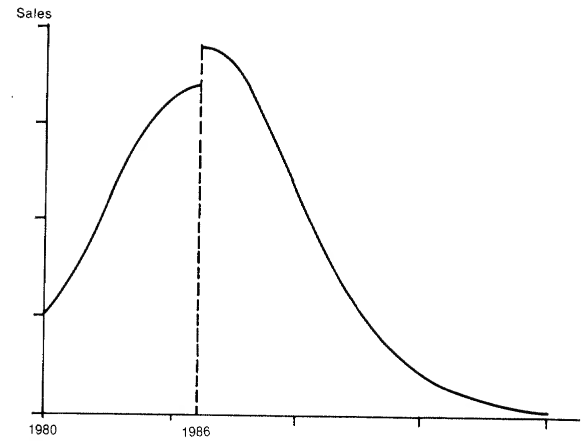


Figure 22–5. Sales Forecast, 1980–1986, with Restraints on the Model

uct by 10 percent. Again the sales projection prior to 1986 is given by figure 22–4, but from 1986 on we have some new conditions: p^* remains at 0.66; the post-1986 number of effective users N_0 is still the original 1980 figure, 5,967, plus the 47,000 units sold between 1980 and 1986, or $N_0 = 52{,}976$. But the remaining market size is no longer 53,000 units. The new post-1986 market size is 1.10 times 53,000, or 58,300 units. Restarting our model in 1986 using these market parameters yields the forecast shown in figure 22–5. The market size increase in 1986 produces a life cycle with a discontinuity in 1986.

Suppose, however, we fear that in 1986, as we introduce the improved product, a major competitor will at the same time introduce a directly competitive product with several major feature improvements and at a price below our own. Clearly we would expect the post-1986 market of 58,300 units to be negatively impacted. Again basic market research and analysis would be necessary to develop an estimate of the exact impact on the post-1986 market size. Assume that such an analysis shows that the post-1986 market is reduced to a mere 15,000 units. Our new-product forecast would again coincide with that in figure 22–4 up to 1986 and then be generated by the parameters $p^* = 0.66$, $N_0 = 52{,}976$, and $N = 15{,}000$ from 1986 on. This forecast is shown in figure 22–6.

It is, of course, possible to stop and restart the model as often as one wishes. Thus, one can deal with multiple pricing changes, competitive introductions, and other elements which alter the remaining market size for the product with the passage of time. For example, if one wishes to incorporate yearly growth in the market size (due to the growth in households or business firms), it is a simple process to reinitiate the model year, adding the

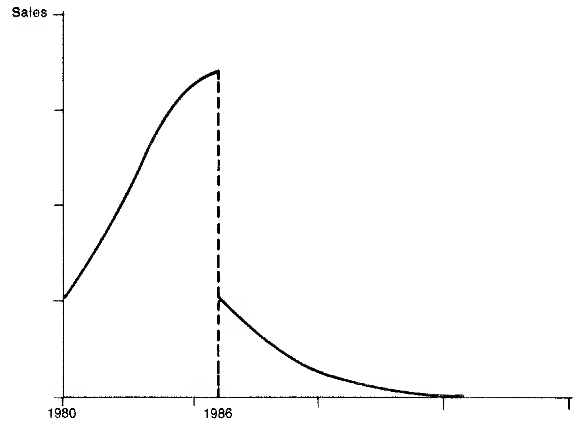


Figure 22–6. Product Forecast, 1980–1986, with Competition

previous year's sales to N_0 and suitably increasing the remaining market size to reflect the current year's growth.

Another major factor which results in the time change of market size involves the inevitable replacement phenomenon. Eventually, the originally purchased product will wear out, and a new replacement purchase will be made. In a stable marketplace this would result eventually in a nearly constant annual-sales situation. The present model deals only with first purchase and thus always predicts that sales will eventually slide asymptotically to zero.

Figure 22–7 shows 8 years' sales for a consumer product which has been fitted with our basic model. Sales are in thousands, and so we are estimating a basic market size of 626,920 units. The forecast out to 1986 is, however, quite depressing, as well as unreasonable. The problem is, of course, that the basic equipment has an average life of about 8 years, and the repurchase phenomenon is not in the basic model. While there are some very complex ways of introducing the repurchase phenomenon into the model, we discuss one of the simpler ones.

The simplest method of approximating the repurchase phenomenon is to go ahead and produce the first-purchase forecast of figure 22–7 shown in table 22–1. These sales show a decline after 1975. Then, as a very crude approximation, simply assume that the equipment wears out and is replaced in exactly 8 years (the average life). This produces a repeat of the original purchase curve shifted 8 years, as shown in table 22–1. Now one adds the two sales columns to produce the overall market sales forecast. The resulting forecast is shown in figure 22–8. Combinations of incremental market growth and repurchase growth can also be created.

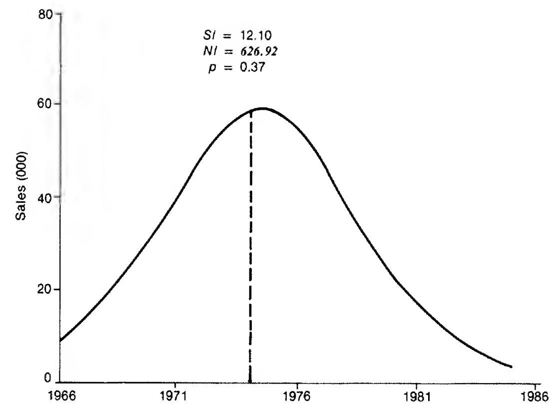


Figure 22-7. Consumer Product Sales Fitted to the Basic Model

Table 22-1
First-Purchase Forecasts

Year	*Original Purhase*	*Replacement Purchase*	*Total*
1967	12.1		12.1
1968	16.7		16.7
1969	22.6		22.6
1970	29.9		29.9
1971	38.2		38.2
1972	46.6		46.6
1973	54.1		54.1
1974	59.1		59.1
1975	60.4	12.1	72.5
1976	57.7	16.7	74.4
1977	51.7	22.6	74.3
1978	43.7	29.9	73.6
1979	35.2	38.2	73.4
1980	27.2	46.6	73.8
1981	20.4	54.1	74.5

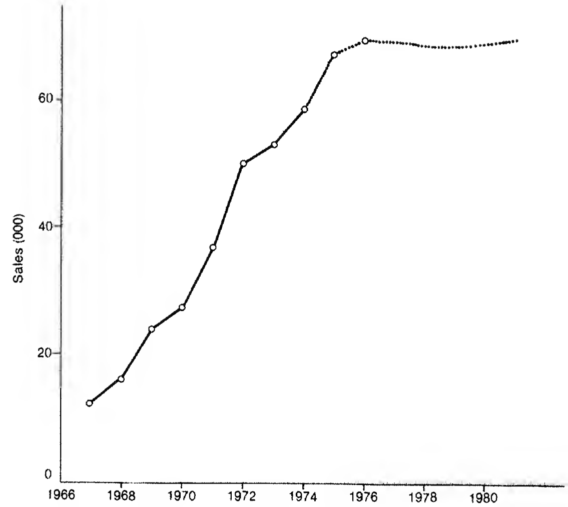


Figure 22-8. Overall Market Sales Forecast

Implementation Enhancements

These model structures have been of important value to market management in several large organizations. They have assisted marketing managers in developing key marketing and sales policies with regard to the levels, placement, and techniques of advertising and promotion, the level of price to adopt, and the markets to sell it. Furthermore, they have provided keen insight into the level of service and production capacity to develop for each product. A majority of the implemented uses of this given model structure have been in cases where no previous sales data existed. Thus, these models provided one of the major decision-making aids to corporate management in developing sound and logically based policies and plans with regard to these products.

The usefulness of the given diffusion-model structure can be enhanced significantly by segmenting the model structure into distinct market segments. Then for each market segment, product-sales forecasts can be developed which provide marketing management with a disaggregated view of the overall marketplace. Such information enhances management's ability to develop marketing plans and strategies which mesh with distinct characteristics of each market segment.

Further possible enhancements to the model include using a range of parameters for market size, the rate parameter, and the effective number of

prior users at the time of introduction. These values can reflect the type of methodology typically associated with that of time estimation in project planning. That is, by using optimistic, most likely, and pessimistic values for each of the parameters, expected-value and variability measures for the sales of the product can be developed. Thus, marketing management is provided with not only a single set of estimates but also a range of probable sales values. These enhancements can be further extended to include the use of policy simulations with various forms of scenarios.

The type of scenarios that will be considered for simulation analysis of the diffusion model will be linked to either potential economic changes (that is, higher or lower rates of inflation) or marketing changes (that is, greater advertising expenditures or differing unit costs).

The economic environment plays a significant role in the sales of the product. Purchasing power in the consumer market is related directly and indirectly to items such as real income growth, the rate of inflation, changing patterns, and the increased cost of energy. Other related items to be considered include the accelerating pace of technological change and the components of technological change. The specific nature of the proposed scenarios can be provided through both Delphic methods and quantitatively based methodologies.

In regard to the marketing environment, we will be concerned with the controllable factors, particularly marketing programs, developed by the sellers. The level of product sales will definitely be affected by promotion, product improvements, and the cost of the product. The factors which are controllable by the firm within limits will also be developed into potential scenarios for evaluation by the simulation modeling process.

References

Bass, F.M. 1969. "A New Product Growth Model for Consumer Durables." *Management Science* 15 (January):215–227.

Davies, S. 1977. *The Diffusion of Process Innovations.* London: Cambridge University Press.

Draper, N.R., and Smith, H. 1967. *Applied Regression Analysis.* New York: John Wiley & Sons.

Geurts, M., and Reinmuth, J.E. 1980. "New Product Sales Forecasting without Post Sales Dates." *European Journal of Operational Research* 4:84–90.

Lawton, S.B., and Lawton, W.H. 1979. "An Autocatalytic Model for the Diffusion of Educational Innovations." *Educational Administration Quarterly* 15 (1):19–53.

Mahajan, V., and Muller, E. 1979. "Innovation Diffusion and New Product Growth Models in Marketing." *Journal of Marketing* 43 (Fall): 55–68.

Midgley, D.F. 1977. *Innovation and New Product Marketing*. New York: Halstead Press.

Index

Index

List of Contributors

Frank M. Bass is Loeb Distinguished Professor of Management, Krannert Graduate School of Management, Purdue University, West Lafayette, Indiana.

Robert Blattberg is a professor of marketing and statistics, University of Chicago.

Philip C. Burger is chairman, Management Systems Group, and associate professor at the State University of New York at Binghamton.

J. Douglas Carroll is a member of the technical staff, Department of Special Communications Research, Bell Laboratories, Murray Hill, New Jersey.

Henry J. Claycamp is vice president, Corporate Planning, International Harvester Company, Chicago.

B.J.K. Collins is manager of the Marketing Interpretation Department of Attwood Statistics, United Kingdom.

Joe A. Dodson is vice president, NW Ayer ABH International, New York.

David Flath is an assistant professor, Department of Economics and Business, North Carolina State University, Raleigh.

John Golanty is president, Golanty and Blattberg, Inc., Chicago.

Paul E. Green is S.S. Kresge Professor of Marketing, The Wharton School, University of Pennsylvania, Philadelphia.

Howard Gundee is vice president, Elrick and Lavidge, Inc., Chicago.

Manohar U. Kalwani is an associate professor of management, Krannert School of Management, Purdue University, West Lafayette, Indiana.

Robert Lavidge is president, Elrick and Lavidge, Inc., Chicago.

Kenneth D. Lawrence, Ph.D., is group manager, Planning and Financial Analysis, AT&T Long Lines, Bedminster, New Jersey.

William H. Lawton, Ph.D., is director of marketing research, Far East and Africa Division, Eastman Kodak Co., Rochester, New York.

E.W. Leonard is an assistant professor, School of Business, Emory University, Atlanta, Georgia.

Joel Levine is vice president, Marketing Research, The Pillsbury Company, Minneapolis, Minnesota.

Lucien E. Liddy is retired from NW Ayer, New York.

David F. Midgley is a senior lecturer, University of New South Wales, Australia.

Eitan Muller is a lecturer of marketing, The Hebrew University of Jerusalem.

Chakravarthi Narasimhan is an assistant professor of marketing, University of Chicago.

J.H. Parfitt is chairman, Mass Observation U.K., Ltd., London, England.

Vithala R. Rao is a professor of marketing, Cornell University, Ithaca, New York.

Patrick J. Robinson is president, Robinson Associates, Inc.; adjunct professor of marketing and business administration, Drexel University; and senior lecturer, The Wharton School, University of Pennsylvania.

Subrata K. Sen is an associate professor of business administration, University of Rochester, Rochester, New York.

Allan D. Shocker is an associate professor of business administration, Graduate School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania.

Alvin J. Silk is a professor of management science, Sloan School of Management, Massachusetts Institute of Technology, Cambridge.

V. Srinivasan is a professor of marketing and management science, Graduate School of Business, Stanford University, Stanford, California.

Edward M. Tauber is professor and chairman, Department of Marketing, University of Southern California, Los Angeles.

Glen L. Urban is a professor of management science, Sloan School of Management, Massachusetts Institute of Technology, Cambridge.

Yankelovich, Skelly and White, Inc., 575 Madison Avenue, New York, New York.

About the Editors

Yoram Wind is a professor of marketing and director of the Wharton Center for International Management Studies at The Wharton School, University of Pennsylvania. He joined the Wharton staff in 1967 after receiving the Ph.D. degree from Stanford University. Dr. Wind is the current editor of *Journal of Marketing* and an area editor of the new *Journal of Marketing Science.* He serves on the editorial boards of all the major marketing journals and as an occasional reviewer to numerous other professional journals. He is the consulting editor of Addison-Wesley's marketing series.

Dr. Wind is the recipient of a number of research awards. He has served as a marketing consultant and has conducted research studies for over fifty companies. In addition, he is a frequent lecturer in faculty seminars at numerous universities and various executive programs both in the United States and abroad.

Dr. Wind is the author of *Product Policy* (forthcoming) and coauthor of several books, including *Advertising Measurement and Decision Making, Market Segmentation, Organizational Buying Behavior,* and *Multi-Attribute Decisions in Marketing.* He has published numerous articles and monographs, with special emphasis on marketing research, product and marketing strategy, and consumer and industrial buyer behavior.

Dr. Wind is an active member of a number of professional associations. He is a past chairman of the Institute of Management Science, College of Marketing, and the policy board of *Journal of Consumer Research.* He is a current member of the American Marketing Association publication board, the board of directors of the Philadelphia Chapter of the American Marketing Association, and the chairman of the Parlin Board of Governors.

Vijay Mahajan is an associate professor of marketing at The Wharton School, University of Pennsylvania. He received the B.S. in chemical engineering from the Indian Institute of Technology, Kanpur, India, and the M.S. in chemical engineering and Ph.D. in management from the University of Texas at Austin. He has published extensively in journals such as *Management Science, Journal of Marketing Research, Journal of Marketing,* and *Harvard Business Review.* Professor Mahajan is coeditor of *Systems Analysis in Health Care* and serves on the editorial boards of *Journal of Marketing* and *Journal of Retailing.* He has consulted for both industry and government and has served on the faculties of Ohio State University and the State University of New York at Buffalo.

Richard N. Cardozo is professor of marketing and director of the Case Development Center at the Graduate School of Business Administration, University of Minnesota. He has served as chairman of the Department of Marketing and director of the university's Center for Experimental Studies in Business. His teaching and research activities concern product policy, marketing strategy and planning, and organizational buying behavior.

Professor Cardozo is the author of *Product Policy: Cases and Concepts,* and a coauthor of *Problems in Marketing*. He has contributed articles to journals such as *Journal of Marketing, Journal of Marketing Research, Journal of Advertising Research, Journal of Retailing, Journal of Social Psychology,* and *Industrial Marketing Management.*

Professor Cardozo consults extensively in his field and participates in numerous executive-education programs. He has been a director of the American Marketing Association and the Product Development and Management Association.

Professor Cardozo received the Ph.D. from the University of Minnesota, the M.B.A. from the Harvard Business School, and the A.B. from Carleton College. He has received a Fulbright grant to study at the London School of Economics and lecture at the Hebrew University of Jerusalem. Before coming to the University of Minnesota, he taught at the Harvard Business School.